

Reconstruction/Non-reconstruction Thresholds for Colourings of General Galton-Watson Trees.

Charilaos Efthymiou

College of Computing, Georgia Institute of Technology, Atlanta, USA
efthymiou@gmail.com

Abstract

The broadcasting models on trees arise in many contexts such as discrete mathematics, biology, information theory, statistical physics and computer science. In this work, we consider the k -colouring model. A basic question here is whether the root's assignment affects the distribution of the colourings at the vertices at distance h from the root. This is the so-called *reconstruction problem*. For the case where the underlying tree is d -ary it is well known that $d/\ln d$ is the *reconstruction threshold*. That is, for $k = (1 + \epsilon)d/\ln d$ we have non-reconstruction while for $k = (1 - \epsilon)d/\ln d$ we have reconstruction.

Here, we consider the largely unstudied case where the underlying tree is chosen according to a predefined distribution. In particular, our focus is on the well-known Galton-Watson trees. This model arises naturally in many contexts, e.g. the theory of spin-glasses and its applications on random Constraint Satisfaction Problems (rCSP). The aforementioned study focuses on Galton-Watson trees with offspring distribution $\mathcal{B}(n, d/n)$, i.e. the binomial with parameters n and d/n , where d is fixed. Here we consider a *broader* version of the problem, as we assume *general offspring distribution*, which includes $\mathcal{B}(n, d/n)$ as a special case.

Our approach relates the corresponding bounds for (non)reconstruction to certain *concentration properties* of the offspring distribution. This allows to derive reconstruction thresholds for a very wide family of offspring distributions, which includes $\mathcal{B}(n, d/n)$. A very interesting corollary is that for distributions with expected offspring d , we get reconstruction threshold $d/\ln d$ under *weaker concentration* conditions than what we have in $\mathcal{B}(n, d/n)$.

Furthermore, our reconstruction threshold for the random colorings of Galton-Watson with offspring $\mathcal{B}(n, d/n)$, implies the reconstruction threshold for the random colourings of $G(n, d/n)$.

1 Introduction

The broadcasting models on trees and the closely related reconstruction problem are studied in statistical physics, biology, communication theory, e.g. see [9, 26, 14]. Our work is motivated from the study of *random Constraint Satisfaction Problems* (rCSP) such as random graph colouring, random k -SAT etc. This is mainly because the models on random trees capture some of the most fundamental properties of the corresponding models on random (hyper)graphs, [8, 15, 24].

The most fundamental problem in the study of broadcasting models is to determine the reconstruction/non-reconstruction threshold. I.e. whether the configuration of the root biases the distribution of the configuration of distant vertices. The transition from non-reconstruction to reconstruction can be achieved by adjusting appropriately the parameters of the model. Typically, this transition exhibits a *threshold behaviour*.

So far, the main focus of the study was to determine the precise location of this threshold for various models when the underlying graph is a fixed tree, mostly regular. In a lot of applications, e.g. phylogeny reconstruction, rCSP, usually the underlying tree is random. Motivated by such problems, in this work we study the reconstruction problem for the colouring model when the underlying tree is chosen

according to some predefined probability distribution. In particular, we consider *Galton-Watson* trees (GW-trees) with some *general* offspring distribution.

The main technical challenge is to deal with is the so-called “effect of high degrees”. That is, we expect to have vertices in the tree which are of degree much higher than the expected offspring. The deviation from the expected degree is so large that expressing the (non)reconstruction bounds in terms of maximum degree leads to highly suboptimal results. Similar challenges appear in problems in random graphs $G(n, d/n)$ e.g. sampling colourings [11, 12, 13, 31].

It is a folklore conjecture that when the offspring distribution is “reasonably” concentrated about its expectation, then the reconstruction threshold can be expressed in terms of the expected offspring of the underlying tree. Somehow, the concentration makes the high degree vertices sufficiently rare, such that their effect on the phenomenon is negligible. Our aim is to make the intuitive base of this relation *rigorous* by just adopting the most generic assumptions about the offspring distribution.

More specifically, our result summarizes as follows: We provide a concentration criterion for the distributions over the non-negative integers about the expectation. For a GW-tree with offspring distribution that satisfies this criterion, the transition from non-reconstruction to reconstruction exhibits a threshold behaviour at the critical point $d/\ln d$, where d is the expected offspring.

Interestingly, the aforementioned concentration criterion is much weaker than the standard tail bounds we have for many natural distributions, e.g. $\mathcal{B}(n, d/n)$. On the other hand, when the concentration of the offspring distribution is not sufficiently high to provide thresholds, we still get upper and lower bounds for reconstruction and non-reconstruction, respectively. These bounds are expressed in terms of the tails of the offspring distribution.

Concluding, let us remark that the reconstruction threshold we get for the random colourings of GW-tree with offspring $\mathcal{B}(n, d/n)$, allows to compute the corresponding threshold for the random colourings of $G(n, d/n)$ [8, 15, 24]. See Section 2.1 for more discussion.

2 Definitions and Results

For the sake of brevity, we define the colouring model and the reconstruction problem, first, in terms of a fixed complete Δ -ary T of height h , where $\Delta, h > 0$ are integers. Later we will extend these definitions w.r.t. GW trees.

The broadcasting models on a tree T are models where information is sent from the root over the edges to the leaves. For some finite set of spins (colours) $S = \{1, 2, \dots, k\}$, a configuration on T is an element in S^T , i.e. it is an assignment of spins to the vertices of T . The spin of the root r is chosen according to some initial distribution over S . The information propagates along the edges of the tree as follows: There is a $k \times k$ stochastic matrix M such that if the vertex v is assigned spin i , then its child u is assigned spin j with probability $M_{i,j}$. The k -colouring model we consider here corresponds to having M such that

$$M_{i,j} = \begin{cases} \frac{1}{k-1} & \text{for } i \neq j \\ 0 & \text{otherwise.} \end{cases}$$

We let μ be the *uniform distribution* over the k -colourings of T . We also refer to μ as the Gibbs distribution. Fixing the spin (colour assignment) at the root of T , the configuration we get after the process has finished is distributed as in μ conditional the spin of the root.

The reconstruction problem can be cast very naturally in terms of the corresponding Gibbs distribution. More specifically, let $r(T)$ (or r_T) denote the root of the tree T . Also, let $L_h(T)$ be the set of vertices at distance h from the root $r(T)$. Finally, we let μ^i be the distribution μ conditional that the spin at r_T is i . Reconstructibility is defined as follows:

Definition 1 For any $i, j \in S$ let $\|\mu^i - \mu^j\|_{L_h}$ denote the total variation distance of the projections of μ^i and μ^j on L_h . We say that a model is reconstructible on a tree T if there exists $i, j \in S$ for which

$$\lim_{h \rightarrow \infty} \|\mu^i - \mu^j\|_{L_h(T)} > 0.$$

When the above limit is zero for every i, j , then we say that the model has non-reconstruction.

Non-reconstruction implies, also, that *typical* colourings of the vertices at level h of the tree have a vanishing effect on the distribution of the colouring of $r(T)$, as h grows.

For the colouring model on Δ -ary trees it is well-known that the reconstruction threshold is $\Delta / \ln \Delta$, see [2, 27, 29, 30]. That is, for any given fixed $\epsilon > 0$ and sufficiently large Δ , i.e. $\Delta \geq \Delta(\epsilon)$, when $k \geq (1 + \epsilon)\Delta / \ln \Delta$ we have non-reconstruction while for $k \leq (1 - \epsilon)\Delta / \ln \Delta$ we have reconstruction.

Rather than considering a fixed tree, here, we consider a Galton Watson tree (GW-trees) with some *general* offspring distribution. In particular, we let the following:

Definition 2 Let ξ be a distribution over the non negative integers. We let \mathcal{T}_ξ denote a Galton-Watson tree with offspring distribution ξ . Also, given some integer $h > 0$, we let \mathcal{T}_ξ^h denote the restriction of \mathcal{T}_ξ to its first h levels¹.

For the sake of brevity any distribution ξ on the non-negative integers is represented as a stochastic vector. That is, for Z distributed as in ξ it holds that $\Pr[Z = i] = \xi(i)$ (or ξ_i), for any integer $i \geq 0$. The notion of reconstruction/non-reconstruction from Definition 1, extends as follows for Galton-Watson trees:

Definition 3 We say that a model is reconstructible on \mathcal{T}_ξ if there exists $i, j \in S$ for which

$$\lim_{h \rightarrow \infty} \mathbb{E} \|\mu^i - \mu^j\|_{L_h} > 0,$$

where the expectation is w.r.t. the instances of the tree. When the above limit is zero for every $i, j \in S$, then we say that the model has non-reconstruction.

So as to have a threshold behavior for reconstruction, it is natural to have a certain kind of parametrization for the offspring distribution ξ . This parametrization allows to adjust the expectation from low to high. In what follows we assume that we deal with such distribution.

Definition 4 Consider \mathcal{T}_ξ for some offspring distribution ξ with expected offspring d_ξ . For the k -colouring model on \mathcal{T}_ξ we have a reconstruction threshold θ for some function $\theta : \mathbb{R}^+ \rightarrow \mathbb{R}^+$, if the following holds: For any $\alpha > 0$ and $d_\xi > d_\xi(\alpha)$, we have non-reconstruction when $k \geq (1 + \alpha)\theta(d_\xi)$, while we have reconstruction when $k \leq (1 - \alpha)\theta(d_\xi)$.

One of the main results of this work is to show that we have a threshold behaviour for the reconstruction/non-reconstruction transition for the k -colourings of \mathcal{T}_ξ when ξ is *well concentrated*. The notion of well concentration is defined as follows:

Definition 5 A distribution ξ over the positive integers with expectation d_ξ is defined to be “well concentrated” if the following is true: There is an absolute constant $c > 0$ such that for any fixed $\gamma > 0$, $d_\xi > d_\xi(\gamma)$ and any $x \geq (1 + \gamma)d_\xi$ it holds that

$$\sum_{j \geq x} \xi_j \leq x^{-c} \quad \text{and} \quad \sum_{j \leq (1-\gamma)d_\xi} \xi_j \leq (d_\xi)^{-c}. \quad (1)$$

¹In other words, \mathcal{T}_ξ^h is the induced subtree of \mathcal{T}_ξ which contains all the vertices within graph distance h from the root.

The quantity c is independent of the distribution ξ . We do not compute the exact value of c but it is implicit from our derivations.

The following theorem is one of the main results in our work.

Theorem 1 *Let ξ be a well concentrated distribution over the non-negative integers. Then, the colouring model on \mathcal{T}_ξ has reconstruction threshold $d_\xi / \ln d_\xi$, where d_ξ is the expected offspring.*

The above theorem follows as a corollary of a more general and more technical result, Theorem 2. This theorem is more general as it covers non-threshold cases, too. Given Theorem 2, we provide a proof of Theorem 1 in Section 14.

It is not hard to show that $\mathcal{B}(n, d/n)$ is well concentrated. This follows trivially by just using standard Chernoff bounds (e.g. [28]). Then, Theorem 1 implies the following corollary.

Corollary 1 *Consider \mathcal{T}_ξ where ξ is the distribution $\mathcal{B}(n, d/n)$. Then, the colouring model on \mathcal{T}_ξ has reconstruction threshold $d / \ln d$.*

As a matter of fact, it is elementary to verify that $\mathcal{B}(n, d/n)$ is, by no means, the less well concentrated offspring distribution we can have. That is, a distribution with less heavy tails than $\mathcal{B}(n, d/n)$ can be well concentrated.

2.1 From Galton-Watson trees to Random Graphs

The non-reconstruction phenomenon in rCSP seems to be central in algorithmic problems. In particular, it has been related to the *efficiency* of local algorithms which search for satisfying solutions. That is, when we have non-reconstruction, usually there is an efficient (simple) local algorithm which finds satisfying assignments efficiently e.g. [6, 17]. On the other hand, in the reconstruction regime there is no efficient algorithm which finds solutions. For this reason, the transition from non-reconstruction to reconstruction on rCSPs has been attributed the name “algorithmic barrier” for rCSP², e.g. see [1].

The ingenious, however, mathematically non-rigorous *Cavity Method*, introduced by physicists [22, 18], makes very impressive predictions about the most fundamental properties of rCSP. One of the most interesting parts of these predictions involves the Gibbs distribution and its spatial mixing properties, e.g. the reconstruction problem. The Cavity Method predicts that the spatial mixing properties of the Gibbs distribution over the colouring of $G(n, d/n)$ can be studied by means of the Gibbs distribution of the k -colourings over a Galton-Watson tree with offspring distribution $\mathcal{B}(n, d/n)$. That is, choose some vertex v in $G(n, d/n)$ and some fixed radius neighborhood around v . The projection of Gibbs distribution on this neighborhood is, somehow, “similar” to the corresponding Gibbs distribution over the Galton-Watson tree. The above line of arguments, led to conjecture that the colouring model on a random graph $G(n, d/n)$ has the same reconstruction threshold as that of the GW tree with offspring $\mathcal{B}(n, d/n)$.

All the above consideration from Cavity method have been studied on a rigorous basis in [8, 15, 24]. We have a quite accurate picture of the relation between the local projection of Gibbs distribution on $G(n, d/n)$ and the Gibbs distribution on Galton-Watson trees. In particular, we have mathematically rigorous arguments which imply that indeed the reconstruction thresholds for $G(n, d/n)$ and GW-tree coincide as far as the colouring model is concerned³. That is, Corollary 1 implies that, indeed, the reconstruction threshold for the colouring model on $G(n, d/n)$ is $d / \ln d$.

²We should mention that this observation is empirical as there is no corresponding (rigorous) computational hardness result.

³For more details on the convergence between the distribution on the GW-tree and $G(n, d/n)$, see [8].

3 High Level Description

In this section, we give a high level overview of how do we derive upper and lower bounds for reconstruction and non-reconstruction, respectively. Consider an instance of \mathcal{T}_ξ^h for some distribution ξ over the non-negative integers and some integer $h > 0$.

Remark 1 *For a set of vertices Λ in the tree, we use the term “random colouring of Λ ” to indicate the following way of colouring Λ : Take a random colouring of the tree and keep only the colouring of the vertices in Λ . Also, when we refer to “typical colourings of vertex set Λ ”, we imply that they are typical w.r.t. the aforementioned distribution.*

Depending on the tails of ξ we choose appropriate quantities Δ_+ and Δ_- such that $\Delta_- \leq d_\xi \leq \Delta_+$. Given these two quantities we show that we have non-reconstruction for $k \geq (1 + \alpha)\Delta_+ / \ln \Delta_+$ and we have reconstruction for $k \leq (1 - \alpha)\Delta_- / \ln \Delta_-$, for the colouring model on \mathcal{T}_ξ^h , where $\alpha > 0$ is fixed. We show (non)reconstruction by arguing about the structure of \mathcal{T}_ξ^h .

Non Reconstruction. First, we focus on non-reconstruction. Given Δ_+ , we define a set of structural specifications such that if \mathcal{T}_ξ^h satisfies them, then we have non-reconstruction for $k \geq (1 + \alpha)\Delta_+ / \ln \Delta_+$. We should consider Δ_+ to be a parameter for the specifications.

In particular, given Δ_+ , we introduce the notion of *mixing* vertex. Roughly speaking, a vertex $v \in \mathcal{T}_\xi^h$ is mixing if the following is true: A typical k -colouring of the vertices at level h (e.g. Remark 1) does not bias the colouring of v by too much when $k \geq (1 + \alpha)\Delta_+ / \ln \Delta_+$. A vertex is biased if it is forced to choose from a relatively small set of colours. Perhaps a simple example of a vertex u *not* being mixing is when the subtree rooted at u has minimum degree much larger than Δ_+ .

An inductive definition of a mixing vertex, roughly, is as follows: A non leaf vertex v is mixing if the number of its children is at most Δ_+ while no more than $o(\Delta_+)$ of its children are non-mixing vertices. We consider the leaves of the tree to be mixing vertices, by default.

Furthermore, our specifications require that the mixing vertices are *sufficiently many* and *well spread* in the tree. To be more specific, we want the following: For every path from the root of \mathcal{T}_ξ^h to the vertices at level h a sufficiently large fraction of the vertices is mixing. Additionally, we would like that the number of vertices at level h should not deviate significantly from their expectation.

Then, we argue that non-reconstruction holds for the colouring model on any, *arbitrary*, instance of \mathcal{T}_ξ^h which satisfies the aforementioned specifications when $k \geq (1 + \alpha)\Delta_+ / \ln \Delta_+$. The choice of Δ_+ is the smallest possible that guarantees that \mathcal{T}_ξ^h satisfies the structural specifications with probability that tends to 1 as $h \rightarrow \infty$.

For showing non-reconstruction, given a fixed tree of the desired structure, we use an idea introduced in [4]. The authors there show non-reconstruction by upper bounding appropriately the second moment of a quantity called “magnetization of the root”. This approach has turned out to be quite popular for showing non-reconstruction bounds for various models on fixed trees e.g. [2, 30, 3, 4]. Additionally to [4], our approach builds on the very elegant combinatorial formalization from [2], which uses the notion of *unbiasing boundary* to deal with the magnetization of the root.

The approach in [2] shows non-reconstruction by arguing that the typical colourings of the vertices at level h do not bias the colouring of the vertices in the largest part of the underlying (regular) tree. The additional element here is that the trees we consider are highly non-regular. So as to get a similar effect from the colorings at level h , we need to argue about the subtree structure of each vertex in the tree. At this point we use the specification requirement. In other words, the setting we develop here with the mixing vertices somehow allows to apply the idea of unbiasing boundaries to control the magnetization of the root of the non-regular trees we deal with.

Reconstruction As opposed to non-reconstruction, the reconstruction bound is well known in the special case where the offspring distribution is $\mathcal{B}(n, d/n)$, e.g. [23, 29]. Our approach deviates from both [23, 29] in that it applies to GW-trees with a general offspring distributions, while it focuses on the structural properties of the underlying tree, i.e. as we do for the non-reconstruction bound.

We are based on the following observation. Consider some fixed tree T of height h and some integer $k > 0$. Take a random k -colouring of the vertices at level h of that tree. Consider the probability that the colouring at the root of the tree ‘freezes’ by that random k -colouring. The assignment at the root gets frozen when the colouring of the vertices at level h specifies uniquely the colouring at the root. A sufficient condition for reconstruction is that the probability that the colouring of the root gets frozen is bounded away from zero for any $h > 0$. The reconstruction bound for a Δ -ary tree follows exactly from this argument, i.e. for $k \leq (1 - \alpha)\Delta / \ln \Delta$, the colouring of the root freezes with probability bounded away from zero for any h , see [29, 27].

Somehow, the above arguments imply that if \mathcal{T}_ξ^h has a (Δ_-) -ary subtree, with the same root as \mathcal{T}_ξ^h , then we have reconstruction for $k \leq (1 - \alpha)\Delta_- / \ln \Delta_-$. The structural specification we need for reconstruction is that \mathcal{T}_ξ^h has such a subtree with probability that is bounded away from zero for any $h > 0$. Our choice of Δ_- is the largest possible that guarantees exactly this specification for \mathcal{T}_ξ^h .

Remark 2 *To be more precise, for non-reconstruction the subtree of \mathcal{T}_ξ^h we consider is not exactly Δ_- -ary. The number of children for each non-leaf vertex is very close to Δ_- .*

4 Upper and Lower Bounds

We start our analysis by focusing on the upper and the lower bounds for reconstruction and non-reconstruction, respectively. Consider \mathcal{T}_ξ^h and the k -colouring model on this tree. We define appropriate quantities Δ_- and Δ_+ which depend (mainly) on the statistics of the offspring distribution ξ . As far as Δ_+ is concerned, we have the following:

Definition 6 *Consider a distribution ξ over the non negative integers with expectation d_ξ . Given some fixed $\delta \in (0, 1/10)$, we let $\Delta_+ = \Delta_+(\delta) \geq d_\xi$ be the minimum integer such that the following holds: There is $q \in [0, 3/4]$ and $\beta \geq 4$, independent of d_ξ , such that*

$$q \geq \sum_{i > \Delta_+} \xi_i + \Pr \left[\mathcal{B}(\Delta_+, q) \geq (\Delta_+)^{\delta} \right] \quad (2)$$

and

$$\sum_{t > \Delta_+} t \cdot \xi_t \leq \exp(-2\beta \ln d_\xi), \quad \Pr \left[\mathcal{B}(\Delta_+, q) > (\Delta_+)^{\delta} \right] \leq \exp(-2\beta \ln d_\xi). \quad (3)$$

Given ξ we choose Δ_+ as described above. Then we use Δ_+ as a parameter to specify a set of structural specifications for trees (roughly described in Section 3). For any instance of \mathcal{T}_ξ which satisfies these specification we have non-reconstruction for any $k \geq (1 + \alpha)\Delta_+ / \ln \Delta_+$. The relations between Δ_+ and ξ as specified in (2) and (3) are, essentially, a list of requirements which guarantee that Δ_+ is as close to d_ξ as possible while at the same time \mathcal{T}_ξ^h satisfies the necessary structural specifications with probability that tends to 1 as h grows.

To illustrate the intuition behind the relations in Definition 6, perhaps, it worths focusing on (2). As we mentioned before, the specification requires the tree has sufficiently many and well-spread mixing vertices. Then, it is natural to require that the probability of a vertex in \mathcal{T}_ξ^h to be mixing is sufficiently large regardless of the level of the vertex in the tree. The requirement in (2) guarantees that this probability is appropriately bounded.

To be more specific, a vertex v is mixing if the number of its children is at most Δ_+ , while at most Δ^δ of them are allowed to be non-mixing (δ is as in Definition 6). Let q be an upper bound for the probability of each child of v to be non-mixing⁴. Using elementary arguments, we get that the r.h.s. of (2) is an upper bound for v to be non-mixing. Moreover, if (2) holds, then clearly q is an upper bound for v to be non-mixing, too. That is, if some vertex at some level l of the tree is non-mixing with probability at most q , then (2) guarantees that for any vertex at level $l - 1$ the probability of it being non-mixing has the same upper bound q . This implies that regardless of its level at the tree, each vertex v is mixing with probability at least $1 - q$. The range of q we consider in Definition 6 guarantees that the mixing vertices are as specified by the requirements. For further details see Section 11.

As far as Δ_- is concerned, we have the following.

Definition 7 *Let ξ be a distribution over the non negative integers. Given some $\delta \in (0, 1/10)$, we let $\Delta_- = \Delta_-(\delta) \leq d_\xi$ be the maximum integer such that the following holds: There is $g \in [0, 3/4)$ such that*

$$g \geq \sum_{i < \Delta_-} \xi_i + \sum_{i \geq \Delta_-} \xi_i \cdot \Pr \left[\mathcal{B}(i, 1 - g) < (\Delta_-) - (\Delta_-)^\delta \right]. \quad (4)$$

The arguments for reconstruction are based on showing that with sufficiently large probability the following holds for \mathcal{T}_ξ^h : The root of \mathcal{T}_ξ^h has a subtree of height h such that each non leaf vertex has sufficiently many children, e.g. approximately Δ_- many. We will see in Section 13, that the condition in (4) guarantees that the root of \mathcal{T}_ξ^h has such a subtree with probability bounded away from zero, regardless of the height h . Clearly, this is the structural requirement for reconstruction, we described in Section 3.

The following theorem is the main technical result of our work. The trees considered in Theorem 2 do not necessarily have well concentrated offspring distribution ξ .

Theorem 2 *Let some fixed $\alpha > 0$. Consider an instance of \mathcal{T}_ξ^h such that the expected offspring d_ξ is sufficiently large. Set $\delta = \min\{\alpha/2, 1/10\}$, i.e. the variable that specifies both Δ_+ and Δ_- .*

For μ , the Gibbs distribution over the k -colourings of \mathcal{T}_ξ^h the following is true:

non-reconstruction: *For $k = (1 + \alpha)\Delta_+ / \ln \Delta_+$ and any $i, j \in [k]$ it holds that*

$$\mathbb{E} \|\mu^i - \mu^j\|_{L_h} \leq 8k^2 (2\Delta_+)^{-0.45\delta h}.$$

reconstruction: *For $k = (1 - \alpha)\Delta_- / \ln \Delta_-$ there are $i, j \in [k]$ such that*

$$\mathbb{E} \|\mu^i - \mu^j\|_{L_h} \geq \frac{1}{4} \left(1 - \frac{2}{\log k} \right).$$

Both of the expectations above are taken w.r.t. the tree instances.

The proof of Theorem 2 appears in two sections. In Section 5 we present the proof for the non-reconstruction part. In Section 13 we present the proof for the reconstruction part.

Given Theorem 2, it is elementary to show that Theorem 1 holds. I.e. given that the offspring distribution is well concentrated (Definition 5), we to show that Δ_- and Δ_+ are sufficiently close to each other. The derivations are simple and they are presented in full detail in Section 14.

Notation. For any tree T we let $r(T)$ or r_T denote its root. Let $L_h(T)$ denote the set of vertices at graph distance h from $r(T)$. For every vertex $v \in T$, we define \hat{T}_v the subtree of T as follows: Delete the

⁴The probability of a vertex being non-mixing depends only on the subtree rooted at this vertex.

edge between v and its parent in T . Then \tilde{T}_v is the connected component that contains v . We use the convention that $r(\tilde{T}_v) = v$.

We use capital letter of the Latin alphabet to indicate random variables which are colourings of the tree T , e.g. X, Y , etc. We use small letter of the greek alphabet to indicate fixed colourings, e.g. σ, τ , etc. We use the notation σ_Λ or $X(\Lambda)$ to indicate that the vertices in Λ have a colour assignment specified by the colouring σ or X , respectively.

Given a tree T , we let μ denote the Gibbs distribution for its k -colourings. Usually we consider μ under certain boundary conditions, i.e. given some $\Lambda \subset T$, and some k -colouring of T , σ , we need to consider the Gibbs distribution where the vertices in Λ have fixed colouring σ_Λ . For this case we denote the Gibbs distribution μ^{σ_Λ} . For $\Xi \subseteq T$ we let μ_Ξ denote the *marginal* of the Gibbs distribution for the vertices in Ξ . We denote marginals over the vertex set Ξ of a Gibbs distribution with boundary σ_Λ in the natural way, i.e. $\mu_\Xi^{\sigma_\Lambda}$.

5 Proof of Theorem 2 - Non Reconstruction

First, consider a fixed tree T of height h and we let $L = L_h(T)$. From [25] we have that

$$\|\mu^i - \mu\|_{r_T} \leq k \cdot \sum_{\sigma(L) \in [k]^L} \mu_L(\sigma_L) \cdot \|\mu^{\sigma(L)} - \mu\|_{r_T}. \quad (5)$$

Furthermore, from the definition of the total variation distance we have that

$$\begin{aligned} \sum_{\sigma(L) \in [k]^L} \mu_L(\sigma_L) \cdot \|\mu^{\sigma(L)} - \mu\|_{r_T} &= \frac{1}{2} \sum_{\sigma(L) \in [k]^L} \mu_L(\sigma_L) \cdot \sum_{c \in [k]} \left| \mu_{r_T}^{\sigma(L)}(c) - 1/k \right| \\ &= \frac{1}{2} \sum_{c \in [k]} \sum_{\sigma(L) \in [k]^L} \mu_L(\sigma_L) \cdot \left| \mu_{r_T}^{\sigma(L)}(c) - 1/k \right|. \end{aligned} \quad (6)$$

The quantity $\left| \mu_{r(T)}^{\sigma(L)}(c) - 1/k \right|$, is usually called *magnetization of the root $r(T)$* , e.g. see [5]. The inner sum is the average magnetization at the root, w.r.t. boundaries at the set L . We bound this average magnetization by using the following standard result.

Proposition 1 *Consider a fixed tree T of height h and some integer $k > 0$. For every $c \in [k]$ the following is true: Let X be a random k -colouring of T conditional that $X(r_T) = c$. It holds that*

$$\sum_{\sigma(L) \in [k]^L} \mu_L(\sigma(L)) \cdot \left| \mu_{r_T}^{\sigma(L)}(c) - 1/k \right| \leq \sqrt{\frac{1}{k} \cdot \left\| \mu^{X_L}(\cdot) - \mu^{Z_L^q}(\cdot) \right\|_{\{r_T\}}}, \quad (7)$$

where Z^q is random colouring of T conditional that $Z^q(r_T) = q$, where q maximizes the r.h.s. of (7).

Our proof of Proposition 1, which is very similar to the proof of Lemma 1 in [4], appears in Section 12.

The quantity on the r.h.s. of (7) is a deterministic one, i.e. it depends only the tree T , c and k . We let

$$\mathbb{G}_{c,k}(T) = \left\| \mu^{X_L}(\cdot) - \mu^{Z_L^q}(\cdot) \right\|_{\{r_T\}}.$$

Consider \mathcal{T}_ξ^h as in the statement of Theorem 2. The quantity $\mathbb{G}_{c,k}(\mathcal{T}_\xi^h)$ is a random variable. In the light of (6), (5) and Proposition 1, it suffices to show that $\mathbb{E} \left[\mathbb{G}_{c,k}(\mathcal{T}_\xi^h) \right]$ tends to zero with h sufficiently fast, for any $c \in [k]$.

Definition 8 (Mixing Root) Let Δ_+ and δ be as in the statement of Theorem 2. For a tree T of height h , its root is called mixing if the following holds: When $h = 0$, then $r(T)$ is mixing, by default. When $h > 0$, $r(T)$ is mixing if and only if $\deg(r_T) \leq \Delta_+$ and there are at most $(\Delta_+)^{\delta}$ many vertices v children of $r(T)$ such that \tilde{T}_v does not have a mixing root.

Definition 9 Given $\zeta \in [0, 1]$ and some integer $t > 0$, we let $\mathcal{A}_{t,\zeta}$ denote the set of trees T of height at most t such that the following holds: Every path \mathcal{P} of length h from $r(T)$ to $L_t(T)$ contains at least $(1 - \zeta)t$ vertices v such that \tilde{T}_v has a mixing root.

Before presenting our next result, we need to do the following remad. In Definition 6, given ξ and δ , among others the following inequality should hold for Δ_+ ,

$$\sum_{t \geq \Delta_+} t \cdot \xi_t < \exp(-2\beta \ln d_{\xi}),$$

where $\beta \geq 4$. Given Δ_+ and ξ the exact value of the parameter β is already specified. That is, when we define Δ_+ and ξ , the value of β is implicit.

Proposition 2 Assume that the distribution ξ , δ , Δ_+ are as defined in the statement of Theorem 2. Let $\mathcal{C} = \beta \ln d_{\xi}$. Also, let $\zeta \in (0, 1)$ and $\theta = \theta(\zeta) > 1$ be such that $(1 - \zeta)\theta < 1$ and $\beta(1 - \theta) < -1$. Then, for every $h \geq 1$ it holds that

$$\Pr[\mathcal{T}_{\xi}^h \in \mathcal{A}_{h,\zeta}] \geq 1 - \exp[-(1 - \theta(1 - \zeta))\mathcal{C} \cdot h].$$

The proof of Proposition 2 appears in Section 11.

Theorem 3 Let ξ, δ, Δ_+ and α be as in the statement of Theorem 2. Also, let $\zeta \in (0, 1)$ and let the integer $h \geq 1$. For $k = (1 + \alpha)\Delta_+ / \ln \Delta_+$, it holds that

$$\mathbb{E} \left[\mathbb{G}(\mathcal{T}_{\xi}^h) \mid \mathcal{T}_{\xi}^h \in \mathcal{A}_{h,\zeta} \right] \leq \frac{4(2\Delta_+)^{-0.9(3/4 - \zeta)\delta h}}{\Pr[\mathcal{T}_{\xi}^h \in \mathcal{A}_{h,\zeta}]}.$$

The proof of Theorem 3 appears in Section 6.

Set $\zeta = 1/4$, and $\theta = 1.3$, applying Proposition 2 we get that

$$\Pr[\mathcal{T}_{\xi}^h \notin \mathcal{A}_{h,\zeta}] \leq d_{\xi}^{-0.1h}. \quad (8)$$

For the same values of ζ, θ as above, (8) with Theorem 3 gives that

$$\mathbb{E} \left[\mathbb{G}(\mathcal{T}_{\xi}^h) \mid \mathcal{T}_{\xi}^h \in \mathcal{A}_{h,\zeta} \right] \leq 8(2\Delta_+)^{-0.45\delta h}. \quad (9)$$

Since we always have $0 \leq \mathbb{G}(T) \leq 1$, for ζ and θ as above, we get that

$$\mathbb{E} \left[\mathbb{G}(\mathcal{T}_{\xi}^h) \right] \leq \mathbb{E} \left[\mathbb{G}(\mathcal{T}_{\xi}^h) \mid \mathcal{T}_{\xi}^h \in \mathcal{A}_{h,1/4} \right] + \Pr \left[\mathcal{T}_{\xi}^h \notin \mathcal{A}_{h,1/4} \right] \leq 16(2\Delta_+)^{-0.45\delta h},$$

where the last inequality follows from (8) and (9). The theorem follows.

6 Proof of Theorem 3

Consider first the quantity $\mathbb{G}_{c,k}(T)$, for some fixed tree T . Then, it holds that

$$\mathbb{G}_{c,k}(T) = \left\| \mu^{X_L}(\cdot) - \mu^{Z_L^q}(\cdot) \right\|_{r_T}. \quad (10)$$

An important remark from Proposition 1 is that it allows to use any kind of correlation between the X, Z^q . For this reason we assume that (X, Z^q) is distributed as in $\nu_{c,q}^T$. We are going to specify this distribution soon. First we get the following result.

Proposition 3 *Let ξ, δ, Δ_+ and α be as in the statement of Theorem 3. Also let $0 \leq \gamma \leq \delta$. Then for $k = (1 + \alpha)\Delta_+ / \ln \Delta_+$, it hold that*

$$\begin{aligned} \mathbb{E} \left[\mathbb{G}_{c,k} \left(\mathcal{T}_\xi^h \right) \middle| \mathcal{T}_\xi^h \in \mathcal{A}_{h,\zeta} \right] &\leq \frac{1}{\Pr \left[\mathcal{T}_\xi^h \in \mathcal{A}_{h,\zeta} \right]} \left(2 \exp \left(-\frac{1}{8} (\Delta_+)^{\frac{h/4-1}{2} \delta + \frac{7}{8} \frac{\alpha}{1+\alpha}} \right) \cdot \mathbb{E} \left[\left| L_h \left(\mathcal{T}_\xi^h \right) \right| \right] + \right. \\ &\quad \left. + 2(2(\Delta_+)^{-\gamma})^{(3/4-\zeta)h} \cdot \mathbb{E}[H(X_L, Z_L^q)] \right). \end{aligned} \quad (11)$$

For the above proposition we remark the following: On the r.h.s. of (11) the rightmost expectation term is w.r.t. both the joint distribution of X, Z^q and the distribution over the tree \mathcal{T}_ξ^h . The rest expectations are w.r.t. the distributions over trees only, i.e. \mathcal{T}_ξ^h . The proof of Proposition 3 appears in Section 7.

For showing the theorem we bound appropriately the two expectations on the r.h.s. of (11). It is elementary that

$$\mathbb{E} \left[\left| L_h \left(\mathcal{T}_\xi^h \right) \right| \right] = (d_\xi)^h. \quad (12)$$

For bounding $\mathbb{E} [H(X_L, Z_L^q)]$ we need to specify a coupling between the random variables X and Z^q which minimizes their expected Hamming distance. Observe that the expected hamming distance is both w.r.t. the coupling and the randomness of the trees.

The coupling of X and Z^q we use, can be defined inductively as follows: We colour the vertices from the root down to the leaves. For a vertex v whose father w is such that $X(w) = Z^q(w)$ we couple $X(v)$ and $Z^q(v)$ identically, i.e. $X(v) = Z^q(v)$. On the other hand, when $X(w) \neq Z^q(w)$ we set $X(v) = Z^q(v)$ unless $X(v) = Z^q(w)$, then we set $Z^q(v) = X(w)$.

Let w be a vertex in the tree and let u be a child of w . Then, for the coupling above, it holds that

$$\Pr [X(u) \neq Z^q(u) | X(w) \neq Z^q(w)] = k^{-1}.$$

In \mathcal{T}_ξ^h , the expected number of children per (non-leaf) vertex is d_ξ . Then, it is elementary to show that for a disagreeing vertex, the expected number of disagreeing children is $d_\xi/k \leq \frac{\ln \Delta_+}{1+\alpha}$, since $\Delta_+ > d_\xi$. Furthermore, it holds that

$$\mathbb{E}[H(X_L, Y_L)] \leq \left(\frac{\ln \Delta_+}{(1 + \alpha)} \right)^h. \quad (13)$$

Observe that the above expectation is w.r.t. *both* tree instances and random colourings.

The theorem follows be combining (13), (12) and Proposition 3.

7 Proof of Proposition 3

The previous setting allows to use ideas based on the notion of biasing-unbiasing boundary (introduced in [2]) to prove Proposition 3. To be more precise, the definition of biasing non-biasing boundaries we use here is slightly different than that [2], but the approach is similar.

Definition 10 (Non-Biasing Boundary) For $\alpha, \gamma, \delta, \Delta_+$ as in the statement of Proposition 3, we let $k = (1 + \alpha)\Delta_+ / \ln \Delta_+$, and let some integer $t \geq 1$. Consider a tree H of height t such that $r(H)$ is mixing. For a k -colouring of H σ we say that σ_L does not bias the root if the following holds:

- if $t = 1$, then $\sigma(L_t(G))$ uses all but at least $(\Delta_+)^{\gamma}$ many colours.
- if $t > 1$, then the following holds: We let v_1, \dots, v_s are the children of the root of H , where $s \leq \Delta_+$. Also, let $\mathbb{S} \subseteq \{\tilde{H}_{v_1}, \tilde{H}_{v_2}, \dots, \tilde{H}_{v_s}\}$ contain only the subtrees whose roots are mixing. Then, there are at most Δ_+^{δ} many subtrees $\tilde{H}_{v_i} \in \mathbb{S}$ such that $\sigma(L_{t-1}(\tilde{H}_{v_i}))$ biases the root $r(\tilde{H}_{v_i})$.

Also, we let $\mathcal{U}(T)$ denote the set of all boundary conditions on L which are not biasing.

Note the notion of non-biasing boundary condition makes sense only for trees with mixing roots.

Lemma 1 Let γ, α, Δ_+ be as in the statement of Proposition 3. Let $k = (1 + \alpha)\frac{\Delta_+}{\ln \Delta_+}$, also let some integer $t \geq 1$. Consider a fixed tree T of height t and let $L = L_t(T)$. For σ , a k -colouring of T , such that σ_L is biasing for the root of T the following is true: There is at least one $c \in [k]$ such that for X , a random k -colouring of T , it holds that

$$\Pr[X_{r(T)} = c | X_L = \sigma_L] \geq (\Delta_+)^{-\gamma}.$$

The proof of Lemma 1 appears in Section 10.1.

Definition 11 Let $\alpha, \gamma, \delta, \Delta_+, h$ be as in the statement of Proposition 3. Consider a tree T of height h and let $L = L_h(T)$. For every vertex $w \in L$ we define the set of boundaries $\mathcal{U}_w \subseteq [k]^L$ as follows: Let \mathcal{P} denote the path that connects r_T and w and we let

$$\mathcal{M} = \left\{ v \in \mathcal{P} : \text{dist}(r_T, v) \leq \frac{3}{4}h, \tilde{T}_v \text{ has mixing root} \right\}.$$

Then \mathcal{U}_w contains the boundary conditions on L which do not bias the root of any of the subtrees \tilde{T}_v where $v \in \mathcal{M}$.

Proposition 4 Let $\alpha, \gamma, \delta, \Delta_+, h, \zeta$ be as in the statement of Proposition 3. Let some fixed tree $T \in \mathcal{A}_{h, \zeta}$ and let $L = L_h(T)$. Consider σ, τ to be two k -colourings of T such that $H(\sigma_L, \tau_L) = 1$. Furthermore, assume that $\sigma(w) \neq \tau(w)$ for some $w \in L$, while both $\sigma_L, \tau_L \in \mathcal{U}_w$. Then it holds that

$$\|\mu^{\sigma_L} - \mu^{\tau_L}\|_{r(T)} \leq \Delta_{\zeta, h}^* = (2\Delta_+^{-\gamma})^{(3/4-\zeta)h}.$$

The proof of Proposition 4 appears in Section 8.

Proposition 5 Let $\alpha, \gamma, \delta, \Delta_+, h, \zeta$ be as in the statement of Proposition 3. Consider a fixed tree $T \in \mathcal{A}_{h, \zeta}$. Let X be a random k -colouring of T . For $k = (1 + \alpha)\Delta_+ / \ln \Delta_+$ and any $w \in L_h(T)$ it holds that

$$\Pr[X_L \notin \mathcal{U}_w] \leq 2 \exp \left(-\frac{1}{8}(\Delta_+)^{\frac{h/4-1}{2}\delta + \frac{7}{8}\frac{\alpha}{1+\alpha}} \right).$$

The proof of Proposition 5 appears in Section 9.

Proof of Proposition 3: First, consider some fixed tree $T \in \mathcal{A}_{h,\zeta}$ and we let $L = L_h(T)$. Usually we fix a colouring of L and we call it (the colouring) boundary condition. We also use the term “free” boundary to indicate the absence of any boundary condition on L or some of its vertices.

Consider two colourings of the leaves $\sigma(L)$ and $\tau(L)$. We let m be the Hamming distance between $\sigma(L)$ and $\tau(L)$, i.e. $m = H(\sigma_L, \tau_L)$. Let v_1, \dots, v_m be the vertices in L for which σ_L and τ_L disagree. Consider the sequence of boundary conditions $Z_0, \dots, Z_{2m} \in [k]^L$ such that $\sigma_L = Z_1, \tau_L = Z_{2m}$ while the rest of the members are as follows: For $i \leq m$, we get Z_i from Z_{i-1} by substituting the assignment of v_i from $\sigma(v_i)$ to “free”. Also, for $i \geq m$ we get Z_{i+1} from Z_i by substituting $Z(v_{i-m})$ from “free” to $\tau(v_{i-m})$. It is direct that $H(Z_i, Z_{i+1}) = 1$.

It holds that

$$\|\mu^{\sigma_L} - \mu^{\tau_L}\|_{r(T)} \leq \sum_{i=0}^{2m-1} \|\mu^{Z_i} - \mu^{Z_{i+1}}\|_{r(T)}. \quad (14)$$

Also, it is not hard to see that for every $w \in L$ the following is true: if $\sigma_L \in \mathcal{U}_w$, then $Z_i \in \mathcal{U}_w$ for every $i = 1, \dots, m$. Similarly, if $\tau_L \in \mathcal{U}_w$, then $Z_i \in \mathcal{U}_w$ for every $i = m, \dots, 2m$.

Let the event $\mathbb{U}_{v_i}^{\sigma, \tau} = “\sigma_L \notin \mathcal{U}_{v_i} \cup \tau_L \notin \mathcal{U}_{v_i}”$. Then it holds that

$$\|\mu^{Z_i} - \mu^{Z_{i+1}}\|_{r(T)} \leq \mathbb{I}_{\{\mathbb{U}_{v_i}\}} + \left(1 - \mathbb{I}_{\{\mathbb{U}_{v_i}\}}\right) \Delta_{\zeta, h}^*, \quad (15)$$

where $\Delta_{\zeta, h}^*$ is defined in the statement of Proposition 4. In words, the above inequality states the following: if at least one of the σ_L, τ_L are not in \mathcal{U}_{v_i} , then the l.h.s. of (15) is at most 1. On the other hand, if both $\sigma_L, \tau_L \in \mathcal{U}_{v_i}$ then the total variation distance on the l.h.s. can be upper bounded by using Proposition 4.

Plugging (15) into (14) we have that

$$\|\mu^{\sigma_L} - \mu^{\tau_L}\|_{r(T)} \leq 2 \cdot \sum_{v \in L_h(T)} \mathbb{I}_{\{\sigma_v \neq \tau_v\}} \cdot [\mathbb{I}_{\{\mathbb{U}_v\}} + (1 - \mathbb{I}_{\{\mathbb{U}_v\}}) \cdot \Delta_{\zeta, h}^*]. \quad (16)$$

Now, we consider the quantity $\mathbb{G}_{c,k}(T)$, i.e. $\mathbb{G}_{c,k}(T) = \|\mu^{X_L} - \mu^{Z_L^q}\|_{r(T)}$. For bounding $\mathbb{G}_{c,k}(T)$ we are going to use (16). That is

$$\begin{aligned} \mathbb{G}_{c,k}(T) &= \|\mu^{X_L} - \mu^{Z_L^q}\|_{r(T)} \leq \sum_{\sigma_L, \tau_L \in [k]^L} \Pr[X_L = \sigma_L, Z_L^q = \tau_L] \cdot \|\mu^{\sigma_L} - \mu^{\tau_L}\|_{r(T)} \\ &\leq 2 \cdot \sum_{\sigma_L, \tau_L \in [k]^L} \Pr[X_L = \sigma_L, Z_L^q = \tau_L] \cdot \sum_{v \in L_h(T)} \mathbb{I}_{\{\sigma_v \neq \tau_v\}} \cdot \left(\mathbb{I}_{\{\mathbb{U}_v^{\sigma, \tau}\}} + \left(1 - \mathbb{I}_{\{\mathbb{U}_v^{\sigma, \tau}\}}\right) \Delta_{\zeta, h}^*\right) \quad [\text{from (16)}] \\ &\leq 2 \cdot \sum_{v \in L_h(T)} \left(\Pr[X(v) \neq Z^q(v), \mathbb{U}_v^{X_L, Z_L^q}] + \Pr[X(v) \neq Z^q(v)] \cdot \Delta_{\zeta, h}^*\right) \\ &\leq 2 \cdot \sum_{v \in L_h(T)} \Pr[\mathbb{U}_v^{X_L, Z_L^q}] + 2 \cdot \sum_{v \in L_h(T)} \Pr[X(v) \neq Z^q(v)] \cdot \Delta_{\zeta, h}^*. \end{aligned}$$

Due to symmetry it holds that $\Pr[X(L) \notin \mathcal{U}_v] = \Pr[Z^q(L) \notin \mathcal{U}_v]$. Using this observation and a union bound, the above inequality implies that

$$\begin{aligned} \mathbb{G}_{c,k}(T) &\leq 4 \sum_{v \in L} \Pr[X(L) \notin \mathcal{U}_v] + \Delta_{\zeta, h}^* \sum_{v \in L} \Pr[X(v) \neq Z^q(v)] \\ &\leq 2 \exp\left(-\frac{1}{8}(\Delta_+)^{\frac{h/4-1}{2}\delta + \frac{7}{8}\frac{\alpha}{1+\alpha}}\right) \cdot |L_h(T)| + 2\Delta_{\zeta, h}^* \cdot \mathbb{E}_{\nu_{c,q}}[H(X_L, Z_L^q)], \end{aligned}$$

where in the last inequality we used Proposition 5 to bound $\Pr[X(L) \notin \mathcal{U}_{\mathcal{P}_v}] \cdot \mathbb{E}_{\nu_{c,q}}[H(X(L), Z^q(L))]$ is the expected Hamming distance between X_L and Z_L^q and depends only on the joint distribution of X, Z^q , which is denoted as $\nu_{c,q}$.

The proposition follows by averaging over \mathcal{T}_ξ^h , conditional that we have a tree in $\mathcal{A}_{h,\zeta}$, that is

$$\begin{aligned} \mathbb{E} \left[\mathbb{G}_{c,k} \left(\mathcal{T}_\xi^h \right) \middle| \mathcal{T}_\xi^h \in \mathcal{A}_{h,\zeta} \right] &\leq \frac{1}{\Pr \left[\mathcal{T}_\xi^h \in \mathcal{A}_{h,\zeta} \right]} \left(2 \exp \left(-\frac{1}{8} (\Delta_+)^{\frac{h/4-1}{2} \delta + \frac{7}{8} \frac{\alpha}{1+\alpha}} \right) \cdot \mathbb{E} \left[\left| L_h \left(\mathcal{T}_\xi^h \right) \right| \right] + \right. \\ &\quad \left. + 2(\Delta_+^{-\gamma})^{(3/4-\zeta)h} \cdot \mathbb{E}[H(X_L, Z_L^q)] \right). \end{aligned}$$

The rightmost expectation term is w.r.t. both $\nu_{c,q}$ and the distribution of random trees \mathcal{T}_ξ^h . In the above derivations we used the following, easy to derive, inequality

$$\mathbb{E} \left[f \left(\mathcal{T}_\xi^h \right) \middle| \mathcal{T}_\xi^h \in \mathcal{A}_{h,\zeta} \right] \leq \mathbb{E} \left[f \left(\mathcal{T}_\xi^h \right) \right] / \Pr \left[\mathcal{T}_\xi^h \in \mathcal{A}_{h,\zeta} \right],$$

where f is any non-negative functions on the support of the distribution \mathcal{T}_ξ^h . The proposition follows. \square

8 Proof of Proposition 4

For showing Proposition 4 we use coupling. The coupling is standard and it has been used in different contexts, e.g. [10, 11].

Not at that we have exactly one disagreement only on some vertex $w \in L$ in the tree T . So as to bound $\|\mu^{\sigma_L} - \mu^{\tau_L}\|_{r(T)}$ we take two k -colourings of T, X and Y distributed as in $\mu^{\sigma_L}, \mu^{\tau_L}$ respectively. We are going to couple X, Y and use the fact that

$$\|\mu^{\sigma_L} - \mu^{\tau_L}\|_{r(T)} \leq \Pr[X(r_T) \neq Y(r_T)]. \quad (17)$$

The coupling of the two random variables is done in a step-wise fashion moving away from the disagreeing vertex w . In particular what is of our interest is the vertices on the path \mathcal{P} that connects w with r_T , i.e. $\mathcal{P} = v_0, v_1, \dots, v_h$ where $v_0 = w$ and $v_h = r_T$. We couple X, Y by considering the pairs $(X(v_i), Y(v_i))$, for $i = 1, \dots, h$.

If for some $j \in [h]$ we have that $X(v_j) = Y(v_j)$, then we can couple the remaining vertices in \mathcal{P} identically, i.e. for every $i > j$ we have $X(v_i) = Y(v_i)$. Clearly this holds due to the fact that the underlying graph is a tree. Once we have $X(v_j) = Y(v_j)$ there is no alternative path for the disagreement to propagate to the pairs $X(v_i), Y(v_i)$ for any $i > j$.

On the other hand, consider the case that $X(v_j) \neq Y(v_j)$, for some $h/4 \leq j \leq h$. We need to bound the probability that $X(v_{j+1}) \neq Y(v_{j+1})$ in the coupling. For this we consider two cases, depending on whether the tree $\tilde{T}_{v_{j+1}}$ has a mixing root or not. We show that it holds that

$$\Pr[X(v_{j+1}) \neq Y(v_{j+1}) | X(v_j) \neq Y(v_j)] \leq \begin{cases} 2\Delta_+^{-\gamma} & \text{if } \tilde{T}_{v_{j+1}} \text{ has mixing root} \\ 1 & \text{otherwise.} \end{cases} \quad (18)$$

Once we show that indeed the above bounds hold, it is a matter of straightforward calculations to show that the proposition. In particular, we use (17) and the trivial bound that

$$\|\mu^{\sigma_L} - \mu^{\tau_L}\|_{r(T)} \leq \Pr[X(r_T) \neq Y(r_T)] \leq \prod_{i=h/4}^h \Pr[X(v_i) \neq Y(v_i) | X(v_{i-1}) \neq Y(v_{i-1})].$$

The probabilities on the r.h.s. are substituted by the bounds we have in (18). The theorem then follows by observing that our assumption that $T \in \mathcal{A}_{h,\zeta}$ implies that among the vertices in $\{v_{h/4}, \dots, v_h\}$ there are at least $(3/4 - \zeta)h$ vertices which are mixing roots at their subtree.

Thus, it remains to show the bound in (18). In particular, it suffices to show the bound regarding the case where the $\tilde{T}_{v_{j+1}}$ has mixing root, as the other one is trivial. For this case assume that $X(v_j) = c, Y(v_j) = q$ for two different $c, q \in [k]$. In this situation we have disagreement between $X(v_{j+1}), Y(v_{j+1})$ if either $X(v_{j+1}) = q$ or $Y(v_{j+1}) = c$ or both. Otherwise, i.e. conditional that $X(v_{j+1}) \neq q$ and $Y(v_{j+1}) \neq c$, there is a coupling such that with probability 1, we have $X(v_{j+1}) = Y(v_{j+1})$. Then it becomes apparent that

$$\begin{aligned} \Pr[X(v_{j+1}) \neq Y(v_{j+1}) | X(v_j) = c, Y(v_j) = q] &\leq \\ &\leq \max\{\Pr[X(v_{j+1}) = q | X(v_j) = c], \Pr[Y(v_{j+1}) = c | Y(v_j) = q]\}. \end{aligned}$$

The result follows almost directly. W.l.o.g. consider the term $\Pr[X(v_{j+1}) = q | X(v_j) = c]$. Clearly there is a $c' \in [k]$ such that

$$\Pr[X(v_{j+1}) = q | X(v_j) = c] \leq \Pr[X(v_{j+1}) = q | X(v_j) = c, X(v_{j+2}) = c'].$$

The above holds because $\Pr[X(v_{j+1}) = q | X(v_j) = c]$ can be written as a convex combination of boundaries on v_{j+2} .

We have assumed that $\tilde{T}_{v_{j+1}}$ has mixing root, while $\sigma_L \in \mathcal{U}_w$. Then it is elementary to verify that $\Pr[X(v_{j+1}) = q | X(v_j) = c, X(v_{j+2}) = c'] \leq 2\Delta_+^{-\gamma}$. Essentially, this bound follows by using arguments very similar to those for Lemma 1. We omit the derivations. The proposition follows. \square

9 Proof of Proposition 5

So as to show Proposition 5 we use the following result.

Proposition 6 *Let $\alpha, \gamma, \delta, \Delta_+, \zeta$ be as in the statement of Proposition 5. Let $k = (1 + \alpha)\Delta_+/\ln \Delta_+$. Consider some tree H , of height $t > 0$, which has mixing root. For Z , a random k -colouring of H , the following is true*

$$\Pr[Z_{L_h(H)} \notin \mathcal{U}(H)] \leq \exp\left(-\frac{1}{8}(\Delta_+)^{\frac{t-1}{2}\delta + \frac{7}{4}\frac{\alpha}{1+\alpha}}\right), \quad (19)$$

we remind the reader that $\mathcal{U}(H)$ denote the set of all boundary conditions which are not biasing root.

The proof of Proposition 6 appears in Section 10.

Proof of Proposition 5: The proposition follows by using Proposition 6 and a simple union bound. In particular, let $L = L_h(T)$. Also, let \mathcal{P} denote the path that connects r_T and $w \in L_h(T)$ while

$$\mathcal{M} = \left\{v \in \mathcal{P} : \text{dist}(r_T, v) \leq \frac{3}{4}h, \tilde{T}_v \text{ has mixing root}\right\}.$$

Clearly, $X_L \notin \mathcal{U}_w$ if for some vertex $u \in \mathcal{M}$, it holds that $X(L \cap \tilde{T}_u) \notin \mathcal{U}(\tilde{T}_u)$, i.e the boundary $X(L \cap \tilde{T}_u)$ biases the root of the subtree \tilde{T}_u . That is,

$$\begin{aligned} \Pr[X(L) \notin \mathcal{U}_w] &= \Pr\left[\bigcup_{u \in \mathcal{M}} X_{L \cap \tilde{T}_u} \notin \mathcal{U}(\tilde{T}_u)\right] \leq \sum_{u \in \mathcal{M}} \Pr[X_{L \cap \tilde{T}_u} \notin \mathcal{U}(\tilde{T}_u)] \quad [\text{union bound}] \\ &\leq \sum_{t=(1/4)h}^h \exp\left(-\frac{1}{8}(\Delta_+)^{\frac{t-1}{2}\delta + \frac{7}{8}\frac{\alpha}{1+\alpha}}\right) \leq 2 \exp\left(-\frac{1}{8}(\Delta_+)^{\frac{h/4-1}{2}\delta + \frac{7}{8}\frac{\alpha}{1+\alpha}}\right), \end{aligned}$$

in the last line, above, we used Proposition 6. The proposition follows. \square

10 Proof of Proposition 6

Since we assumed that the tree H has a mixing root, it holds that $\deg(r_H) = s \leq \Delta_+$. We let v_1, v_2, \dots, v_s denote the children of r_H . We remind the reader that the set $\mathbb{S} \subseteq \{\tilde{H}_{v_1}, \tilde{H}_{v_2}, \dots, \tilde{H}_{v_s}\}$ contain only the subtrees whose roots are mixing.

So as to prove Proposition 6 we need the following result.

Lemma 2 *Let X be a random k -colouring H . For $L_i = L_{h-1}(\tilde{H}_{v_i})$, let B_i denote the event that in \tilde{H}_{v_i} , the boundary $X(L_i)$ does not bias $r(\tilde{H}_{v_i})$. For any $\Gamma \subseteq \{1, \dots, s\}$ it holds that*

$$\Pr[\cap_{i \in \Gamma} \mathcal{B}_i] = \prod_{i \in \Gamma} \Pr[\mathcal{B}_i] = (\Pr[\mathcal{B}_i])^{|\Gamma|}.$$

The proof of this lemma is straightforward so we omit it. Essentially, it follows from the fact that a biasing (resp. non-biasing) boundary condition remains biasing (resp. non-biasing) if we repermute the colour classes. A similar lemma appears in [2].

Proof of Proposition 6: The proof is by induction on $t \geq 1$. The induction basis is $t = 1$. Then, H is one level tree whose root is of degree at most Δ_+ . Let Y denote the number of different colours that do not appear in $X(L_1)$. It holds that

$$\Pr[X_{L_1(H)} \notin \mathcal{U}(H)] \leq \Pr[Y \leq \Delta_+^\gamma]. \quad (20)$$

Observe that $\Pr[Y \leq \Delta_+^\gamma]$ is an increasing function of the degree of $r(H)$. That is, the larger the degree of $r(H)$ the more colours are expected to be used to colour the leaves of H . For this reason, we are going to upper bound the r.h.s. of (20) by assuming that $\deg(r_H) = \Delta_+$, i.e. the maximum degree possible for a mixing root. It holds that

$$\begin{aligned} \mathbb{E}[Y] &= (k-1) \left(1 - \frac{1}{k-1}\right)^{\Delta_+} \geq (k-1) \exp\left(-\frac{\Delta_+}{k-2}\right) \quad [\text{as } 1-x \geq e^{\frac{x}{1-x}} \text{ for } 0 < x < 1/5] \\ &\geq (k-1) \exp\left(-\left(1 - \frac{\alpha}{1+\alpha}\right) \ln \Delta_+ - \frac{\ln \Delta_+}{k-2}\right) \geq (\Delta_+)^{\frac{7}{8} \frac{\alpha}{1+\alpha}}. \end{aligned} \quad (21)$$

Viewing the $k-1$ colours which are available for the leaves of H as bins and each leaf of H as a ball which is thrown to a random bin, Y corresponds to the number of empty bins. It is a standard result that we can apply Chernoff bounds for bounding the tails of Y , e.g. see [28]. Then we get that

$$\Pr[Y < (\Delta_+)^{\gamma}] \leq \Pr[Y \leq \mathbb{E}[Y]/2] \leq \exp(-\mathbb{E}[Y]/8) \leq \exp\left(-(\Delta_+)^{\frac{7}{8} \frac{\alpha}{1+\alpha}}/8\right), \quad [\text{as } \gamma \leq \min\{\alpha/2, 1/10\}]$$

where in the last inequality we use (21). We have proved the basis of our induction.

Assume, now, that (19) is true for every tree of height $t-1$ which has mixing root. It suffices to show that (19) is true for a tree H of height t with a mixing root. For such a tree H let $L = L_t(H)$. Consider also a random k -colouring X for this tree. Let Z , denote the number of subtrees in \mathbb{S} which are biased under the random colouring X_L , i.e. the number of trees $\tilde{H}_{v_i} \in \mathbb{S}$ such that $X(L \cap \tilde{H}_{v_i})$ is biasing for $r(\tilde{H}_{v_i})$. From Lemma 1 we have the following

$$\Pr[X_L \notin \mathcal{U}(H)] \leq \Pr[Z > \Delta_+^\delta]. \quad (22)$$

Let

$$\varrho = \max_{\tilde{H}_v \in \mathbb{S}} \left\{ \Pr[X(L \cap \tilde{H}_v) \notin \mathcal{U}(\tilde{H}_v)] \right\},$$

where for the subtree \tilde{H}_v , the set $\mathcal{U}(\tilde{H}_v)$ contains all the boundary conditions (at level $t-1$) \tilde{H}_v which do not bias the root of $r(\tilde{H}_v)$. From Lemma 2 we conclude that Z is dominated by $\mathcal{B}(\Delta_+, \varrho)$, i.e. the binomial distribution with parameters Δ_+ and ϱ . Due to our assumptions it holds that $\Delta_+^\delta \gg \Delta_+ \cdot \varrho$. We have that

$$\begin{aligned} \Pr[Z > \Delta^\delta] &\leq \sum_{j=\Delta_+^\delta}^{\Delta_+} \binom{\Delta_+}{j} \varrho^j (1-\varrho)^{\Delta_+-j} \leq \Delta_+ \binom{\Delta_+}{\Delta_+^\delta} \varrho^{\Delta_+^\delta} (1-\varrho)^{\Delta_+-\Delta_+^\delta} \\ &\leq \frac{\Delta_+}{(\Delta_+/e)^{\Delta_+^\delta}} (\Delta_+ \varrho)^{\Delta_+^\delta} && [\text{as } \binom{n}{i} \leq (ne/i)^i] \\ &\leq (\Delta_+ \varrho)^{\Delta_+^\delta} && \left[\text{as } \frac{\Delta_+}{(\Delta_+/e)^{\Delta_+^\delta}} < 1 \right] \\ &\leq \left(\Delta_+ \exp \left(-\frac{1}{8} \Delta_+^{\frac{t-2}{2}\delta + \frac{7}{8}\frac{\alpha}{1+\alpha}} \right) \right)^{\Delta_+^\delta} && [\text{by the induction hypothesis}] \\ &\leq \left(\exp \left(-\frac{1}{8} \Delta_+^{\frac{t-3}{2}\delta + \frac{7}{8}\frac{\alpha}{1+\alpha}} \right) \right)^{\Delta_+^\delta} \leq \exp \left(-\frac{1}{8} \Delta_+^{\frac{t-1}{2}\delta + \frac{7}{8}\frac{\alpha}{1+\alpha}} \right). \end{aligned} \quad (23)$$

The proposition follows by plugging (23) into (22). \square

10.1 Proof of Lemma 1

The proof is by induction on the height of the tree t . The case where $t = 1$ follows from Definition 10.

Consider some $t > 1$ and assume that the assertion is true for any tree of height less than t . We are going to show that the assertion is true for trees of height t , as well.

Assume that $\deg(r_H) = s$ for some integer s . Clearly $s \leq \Delta_+$ since we assume that H has a mixing root. We let v_1, \dots, v_s be the children of the root. Also, we let $L_i = L \cap \tilde{H}_{v_i}$, where $L = L_t(H)$. That is L_i denotes the vertices at level $t-1$ of the subtree \tilde{H}_{v_i} .

Let X be a random k -colouring of H such that $X_L = \sigma_L$ also, for $i = 1, \dots, s$, let $X_i = X(\tilde{H}_{v_i})$. A standard recursive argument yields the following relation: For any $c \in [k]$ it holds that

$$\Pr[X(r_H) = c] = \frac{\prod_{i=1}^s \Pr[X_i(v_i) \neq c]}{\sum_{c' \in [k]} \prod_{i=1}^s \Pr[X_i(v_i) \neq c']} \leq \frac{1}{\sum_{c' \in [k]} \prod_{i=1}^s \Pr[X_i(v_i) \neq c']}. \quad (24)$$

We show that $r(H)$ if σ_{L_h} is non-biasing then the denominator in (24) is sufficiently small.

Let $B \subset [k]$ denote the set of colours c for which there is some i such that $\Pr[X_i(v_i) = c] \geq \Delta_+^{-\gamma}$. It is only Δ_+^γ many colours can have increased bias at the root of \tilde{H}_{v_i} since $\sum_{c \in [k]} \Pr[X_i(v_i) = c] = 1$.

We have assumed that there are at most Δ_+^δ trees \tilde{H}_{v_i} whose root is mixing but the boundary biases the colour assignment of the root. Furthermore, there are Δ_+^δ trees \tilde{H}_{v_i} with non-mixing roots. That is, there can be at most $2\Delta_+^\delta$ trees \tilde{H}_{v_i} whose roots are biased, those whose root is biased by the boundary condition and those which have non-mixing root.

Clearly, all the above imply that $|B| \leq 2\Delta_+^{\gamma+\delta}$. Letting $U = [k] \setminus B$, we rewrite (24) as follows:

$$\begin{aligned}
\Pr[X(r_H) = c] &\leq \left(\sum_{c' \in U} \prod_{i=1}^s (1 - \Pr[X_i(v_i) = c']) \right)^{-1} \\
&\leq \left(\sum_{c' \in U} \prod_{i=1}^s \exp \left(-\frac{\Pr[X_i = c']}{1 - \Pr[X_i = c']} \right) \right)^{-1} \quad [\text{as } 1 - x > e^{x/(1-x)} \text{ for } 0 < x < 0.1] \\
&\leq \left(|U| \sum_{c' \in U} \frac{1}{|U|} \exp \left(-\sum_{i=1}^s \frac{\Pr[X_i(v_i) = c']}{1 - \Pr[X_i(v_i) = c']} \right) \right)^{-1} \\
&\leq \left(|U| \prod_{c' \in U} \exp \left(-\frac{1}{|U|} \sum_{i=1}^s \frac{\Pr[X_i(v_i) = c']}{1 - \Pr[X_i(v_i) = c']} \right) \right)^{-1} \quad [\text{arithmetic-geometric mean}] \\
&\leq \left(|U| \exp \left(-\frac{1}{|U|} \sum_{i=1}^s \sum_{c \in U} \frac{\Pr[X_i(v_i) = c]}{1 - \Pr[X_i(v_i) = c]} \right) \right)^{-1} \\
&\leq \left(|U| \exp \left(-\frac{1}{|U|} \sum_{i=1}^s \frac{\Pr[X_i(v_i) \in U]}{1 - \Delta_+^{-\gamma}} \right) \right)^{-1} \quad [\text{as } \Pr[X_i(v_i) = c] < \Delta_+^{-\gamma} \text{ for } c \in U] \\
&\leq \left(|U| \exp \left(-\frac{1}{1 - \Delta_+^{-\gamma}} \frac{s}{|U|} \right) \right)^{-1} \quad [\text{as } \Pr[X_i \in U] \leq 1]
\end{aligned}$$

It is straightforward to show that $|U| \geq k \left(1 - \Delta_+^{\frac{\gamma+\delta-1}{2}} \right) \geq (1 + \frac{9}{10}\alpha) \frac{\Delta_+}{\ln \Delta_+}$, since $\gamma + \delta < 1$. Also it holds that $\frac{1}{1 - \Delta_+^{-\gamma}} \frac{s}{|U|} \leq \frac{\ln \Delta_+}{1 + 4\alpha/5}$, since $s \leq \Delta_+$. Thus, we get that

$$\Pr[X = c] \leq \frac{1}{(1 + \alpha/2) \frac{\Delta_+}{\ln \Delta_+} \Delta_+^{-\frac{1}{1+4\alpha/5}}} \leq \Delta_+^{-\frac{3\alpha/5}{1+4\alpha/5}} < \Delta_+^{-\gamma},$$

as $\gamma = \min\{\alpha/2, 1/10\}$. The lemma follows.

11 Proof of Proposition 2

For $i = (1 - \zeta)h$ we let $Q_{h,i} = \Pr \left[\mathcal{T}_\xi^h \notin \mathcal{A}_{h,\zeta} \right]$. Also, we let $Q_{h,i}^t = \Pr \left[\mathcal{T}_\xi^h \notin \mathcal{A}_{h,\zeta} \mid \deg(r(T_\xi^h)) = t \right]$

Using a simple union bound we get the following: For $t \leq (\Delta_+)^{\delta}$ it holds that

$$Q_{h,i}^t \leq t \cdot Q_{h-1,i-1}. \quad (25)$$

Intuitively, the above is implied by the following: If $\deg(r(T_\xi^h)) \leq (\Delta_+)^{\delta}$, then, regardless of its children, the root $r(T_\xi^h)$ is mixing. Conditional that $\deg(r(T_\xi^h)) \leq (\Delta_+)^{\delta}$ holds, so as to have $\mathcal{T}_\xi^h \notin \mathcal{A}_{h,\zeta}$, there should be a vertex v , child of $r(T_\xi^h)$ such that the following is true: The subtree \tilde{T}_v has a path from its root to its vertices of at level $h - 1$ which contain less than $i - 1$ mixing vertices.

Using similar arguments, for $(\Delta_+)^{\delta} \leq t \leq \Delta_+$, we get the following lemma, whose proof appear in Section 11.1.

Lemma 3 For $(\Delta_+)^{\delta} < t \leq \Delta_+$, it holds that

$$Q_{h,i}^t \leq 2t \left(Q_{h-1,i-1} + Q_{h-1,i} \cdot \Pr \left[\mathcal{B}(\Delta_+, q) \geq (\Delta_+)^{\delta} \right] \right).$$

Finally, using a simple union bound we get that for $t > \Delta_+$ it holds that

$$Q_{h,i}^t \leq t \cdot Q_{h-1,i}. \quad (26)$$

The above follows by a line of arguments similar to those we used for (25) and by noting that if $\deg(r(T_\xi^h)) \geq \Delta_+$, then the root of T_ξ^h is non-mixing.

We are bounding $Q_{h,i}$ by using (25), (26) and Lemma 3. We have that

$$\begin{aligned} Q_{h,i} &= \sum_{t=0}^n Q_{h,i}^t \xi_t \\ &= Q_{h-1,i-1} \cdot \sum_{t=0}^{(\Delta_+)^{\delta}} t \cdot \xi_t + 2Q_{h-1,i-1} \cdot \sum_{t=(\Delta_+)^{\delta}+1}^{\Delta_+} t \cdot \xi_t + \\ &\quad + 2Q_{h-1,i} \cdot \Pr \left[\mathcal{B}(\Delta_+, q) \geq (\Delta_+)^{\delta} \right] \cdot \sum_{t=(\Delta_+)^{\delta}+1}^{\Delta_+} t \cdot \xi_t + Q_{h-1,i} \cdot \sum_{t \geq (\Delta_+)+1} t \cdot \xi_t \\ &\leq 2Q_{h-1,i-1} \sum_{t=0}^{\Delta_+} t \cdot \xi_t + Q_{h-1,i} \left(2\Pr \left[\mathcal{B}(\Delta_+, q) \geq (\Delta_+)^{\delta} \right] \sum_{t=(\Delta_+)^{\delta}}^{\Delta_+} t \cdot \xi_t + \sum_{t \geq (\Delta_+)+1} t \cdot \xi_t \right) \\ &\leq 2d_{\xi} \cdot Q_{h-1,i-1} + Q_{h-1,i} \left(2d_{\xi} \cdot \Pr \left[\mathcal{B}(\Delta_+, q) \geq (\Delta_+)^{\delta} \right] + \sum_{t \geq (\Delta_+)+1} t \cdot \xi_t \right). \end{aligned} \quad (27)$$

The following lemma uses (27) to derive an upper bound on $Q_{h,i}$.

Lemma 4 *Let h, β, \mathcal{C} be as in the statement of Proposition 2. Also, let $\lambda \in (0, 1)$ and $\theta' > 1$ be a fixed numbers such that $\beta(1 - \theta') < -1$ and $\lambda\theta' < 1$. Then for $i = \lambda h$ and $Q_{h,i}$ that satisfy the inequality in (27), it holds that*

$$Q_{h,i} \leq \exp \left[-(1 - \lambda\theta') \cdot \mathcal{C} \cdot h \right]. \quad (28)$$

The proof of Lemma 4 appears in Section 11.2

The proposition follows by using the above lemma and setting $\lambda = (1 - \zeta)$ and $\theta' = \theta$, where ζ and θ are defined in the statement of Proposition 2.

11.1 Proof of Lemma 3

Let q_{h-1} be the probability for each child of $r(\mathcal{T}_\xi^h)$ to be non-mixing. Conditional that $r(\mathcal{T}_\xi^h)$ has degree t , the number of non-mixing children of $r(\mathcal{T}_\xi^h)$ is binomially distributed with parameters, t, q_{h-1} , i.e. $\mathcal{B}(t, q_{h-1})$. Letting $Q_{h,i}^M = \Pr \left[\mathcal{T}_\xi^h \notin \mathcal{A}_{h,\zeta} \mid r(\mathcal{T}_\xi^h) \text{ is mixing} \right]$ and $Q_{h,i}^N = \Pr \left[\mathcal{T}_\xi^h \notin \mathcal{A}_{h,\zeta} \mid r(\mathcal{T}_\xi^h) \text{ is not mixing} \right]$, it holds that

$$\begin{aligned} Q_{h,i}^t &\leq \sum_{j=0}^{(\Delta_+)^{\delta}} \binom{t}{j} q_{h-1}^j (1 - q_{h-1})^{t-j} \left[(t-j)Q_{h-1,i-1}^M + jQ_{h-1,i-1}^N \right] + \\ &\quad + \sum_{j=(\Delta_+)^{\delta}+1}^t \binom{t}{j} q_{h-1}^j (1 - q_{h-1})^{t-j} \left[(t-j)Q_{h-1,i}^M + jQ_{h-1,i}^N \right]. \end{aligned}$$

Using the standard equality that $(t-j)\binom{t}{j} = t\binom{t-1}{j}$, we get that

$$\begin{aligned}
Q_{h,i}^t &\leq t(1-q_{h-1})Q_{h-1,i-1}^M \sum_{j=0}^{(\Delta_+)^{\delta}} \binom{t-1}{j} q_{h-1}^j (1-q_{h-1})^{t-1-j} \\
&\quad + tq_{h-1}Q_{h-1,i-1}^N \sum_{j=1}^{(\Delta_+)^{\delta}} \binom{t-1}{j-1} q_{h-1}^{j-1} (1-q_{h-1})^{t-j} \\
&\quad + t(1-q_{h-1})Q_{h-1,i}^M \sum_{j=(\Delta_+)^{\delta}+1}^{t-1} \binom{t-1}{j} q_{h-1}^j (1-q_{h-1})^{t-1-j} \\
&\quad + tq_{h-1}Q_{h-1,i}^N \sum_{j=(\Delta_+)^{\delta}+1}^t \binom{t-1}{j-1} q_{h-1}^{j-1} (1-q_{h-1})^{t-j}.
\end{aligned}$$

It is not hard to see that for any h, i it holds that $q_h Q_{h,i}^N \leq Q_{h,i}$ and $(1-q_h)Q_{h,i}^M \leq Q_{h,i}$. Using these two inequalities we get that

$$\begin{aligned}
Q_{h,i}^t &\leq tQ_{h-1,i-1} \left(\Pr \left[\mathcal{B}(t-1, q_{h-1}) \leq (\Delta_+)^{\delta} \right] + \Pr \left[\mathcal{B}(t-1, q_{h-1}) \leq (\Delta_+)^{\delta} - 1 \right] \right) \\
&\quad + tQ_{h-1,i} \left(\Pr \left[\mathcal{B}(t-1, q_{h-1}) \geq (\Delta_+)^{\delta} + 1 \right] + \Pr \left[\mathcal{B}(t-1, q_{h-1}) \geq (\Delta_+)^{\delta} \right] \right) \\
&\leq 2tQ_{h-1,i-1} + 2tQ_{h-1,i} \Pr \left[\mathcal{B}(t-1, q_{h-1}) \geq (\Delta_+)^{\delta} \right].
\end{aligned} \tag{29}$$

Note that that $\Pr \left[\mathcal{B}(t-1, q_{h-1}) \geq (\Delta_+)^{\delta} \right]$ is increasing with t . That is, for $t \leq \Delta_+$ it holds that

$$\Pr \left[\mathcal{B}(t-1, q_{h-1}) \geq (\Delta_+)^{\delta} \right] \leq \Pr \left[\mathcal{B}(\Delta_+, q_{h-1}) \geq (\Delta_+)^{\delta} \right]. \tag{30}$$

At this point we need to observe that the quantity q , defined in Definition 6, is an upper bound for q_h , for every h . This follows by an inductive argument, i.e. induction on h the number of levels of \mathcal{T}_{ξ}^h .

Clearly, for $h = 0$, the assertion is true. The tree with zero levels consists of only one vertex, which is a leaf. By default the leaves are mixing vertices, i.e. the probability of a leaf to be non-mixing is zero. Since $q \in [0, 3/4]$, q is an upper bound for the vertex to be non-mixing.

Given some $h > 0$, assume that the assertion is true for $\mathcal{T}_{\xi}^{h'}$, for any $h' \leq h$. We are going to show that this is true for \mathcal{T}_{ξ}^h . Let \mathbf{N} be the number of non-mixing children of the root of \mathcal{T}_{ξ}^h . It holds that

$$\Pr[r(\mathcal{T}_{\xi}^h) \text{ is non-mixing}] \leq \Pr[\deg(r(\mathcal{T}_{\xi}^h)) > \Delta_+] + \Pr[\mathbf{N} > (\Delta_+)^{\delta} | \deg(r(\mathcal{T}_{\xi}^h)) \leq \Delta_+].$$

Given that $\deg(r(\mathcal{T}_{\xi}^h)) = D$, for some integer $D \geq 0$, \mathbf{N} is a binomial variable with parameters D, q_{h-1} . Due to our induction hypothesis it holds that $q_{h-1} < q$. Since we have conditioned that $D < \Delta_+$, it is clear that \mathbf{N} is dominated by a binomial variable with parameters Δ_+, q , that is

$$\begin{aligned}
\Pr[r(\mathcal{T}_{\xi}^h) \text{ is non-mixing}] &\leq \Pr[\deg(r(\mathcal{T}_{\xi}^h)) > \Delta_+] + \Pr[\mathcal{B}(\Delta_+, q) > (\Delta_+)^{\delta}] \\
&\leq \sum_{i \geq \Delta_+} \xi_i + \Pr[\mathcal{B}(\Delta_+, q) > (\Delta_+)^{\delta}] \leq q,
\end{aligned}$$

where the last inequality follows from the definition of q , i.e. in Definition 6. The above inequality with (30) imply that

$$\Pr \left[\mathcal{B}(\Delta_+, q_{h-1}) \geq (\Delta_+)^{\delta} \right] \leq \Pr \left[\mathcal{B}(\Delta_+, q) \geq (\Delta_+)^{\delta} \right],$$

as $\mathcal{B}(\Delta_+, q_{h-1})$ is stochastically dominated by $\mathcal{B}(\Delta_+, q)$, since, $q_{h-1} \leq q$, for any h .

The lemma follows by plugging the above inequality into (29).

11.2 Proof of Lemma 4

We are going to use induction to prove the lemma. First we are going to show that if (28) is true for some $h > 1$ then it is also true for $h + 1$. Let $\lambda = \frac{i}{h}$, $\lambda^- = \frac{i-1}{h-1}$ and $\lambda^+ = \frac{i}{h-1}$. We rewrite (27) in terms of λ , λ^+ and λ^- as follows:

$$Q_{\{h, \lambda h\}} \leq 2d \cdot Q_{\{h-1, \lambda^-(h-1)\}} + Q_{\{h-1, \lambda^+(h-1)\}} \left(2d \Pr [\mathcal{B}(\Delta_+, q) \geq (\Delta_+)^{\delta}] + \sum_{t \geq (\Delta_+)+1} t \cdot \xi_t \right). \quad (31)$$

Using the induction hypothesis and noting that $\lambda^- = \lambda - \frac{1-\lambda}{h-1}$ we have that

$$\begin{aligned} Q_{\{h-1, \lambda^-(h-1)\}} &\leq \exp [-(1 - \theta \lambda^-)(h-1)\mathcal{C}] \\ &\leq \exp \left[- \left(1 - \theta' \left(\lambda - \frac{1-\lambda}{h-1} \right) \right) (h-1)\mathcal{C} \right] \\ &\leq \exp [-(1 - \theta' \lambda)(h-1)\mathcal{C}] \cdot \exp [-\theta' (1 - \lambda) \mathcal{C}] \\ &\leq \exp [-(1 - \theta' \lambda) h \mathcal{C}] \cdot \exp [(1 - \theta') \mathcal{C}]. \end{aligned}$$

As far as $Q_{\{h-1, i\}}$ is regarded, we use the fact that $\lambda^+ = \lambda + \frac{\lambda}{h-1}$ and we get that

$$\begin{aligned} Q_{\{h-1, \lambda^+(h-1)\}} &\leq \exp [-(1 - \theta' \lambda^+)(h-1)\mathcal{C}] \\ &\leq \exp \left[- \left(1 - \theta' \lambda - \frac{\theta' \lambda}{h-1} \right) (h-1)\mathcal{C} \right] \\ &\leq \exp [-(1 - \theta' \lambda)(h-1)\mathcal{C}] \cdot \exp [\theta' \lambda \mathcal{C}] \\ &\leq \exp [-(1 - \theta' \lambda) h \mathcal{C}] \exp [\mathcal{C}]. \end{aligned}$$

Substituting the bounds for $Q_{\{h-1, i-1\}}$, $Q_{\{h-1, i\}}$ above into (31) we get that

$$\begin{aligned} Q_{\{h, \lambda h\}} &\leq \exp [-(1 - \theta' \lambda) h \mathcal{C}] \times \\ &\quad \times \left(2d \cdot \exp [(1 - \theta') \mathcal{C}] + \exp (\mathcal{C}) \left(2d \Pr [\mathcal{B}(\Delta_+, q) \geq (\Delta_+)^{\delta}] + \sum_{t \geq (\Delta_+)+1} t \cdot \xi_t \right) \right). \end{aligned}$$

From to our assumption that $\beta(1 - \theta') < -1$ it is direct that

$$2d \cdot \exp [(1 - \theta') \mathcal{C}] = 2d^{1+\beta(1-\theta')} \leq 1/5.$$

Also due to our assumptions about Δ_+ , δ we get that

$$\exp (\mathcal{C}) \left(2d \Pr [\mathcal{B}(\Delta_+, q) \geq (\Delta_+)^{\delta}] + \sum_{t \geq \Delta_++1} t \cdot \xi_t \right) \leq \frac{2}{5}.$$

Using the two bounds above (32) writes as follows:

$$Q_{\{h, \lambda h\}} \leq \exp [-(1 - \theta' \cdot \lambda) h \mathcal{C}].$$

It remains to show the base of the induction, i.e the case $h = 1$. Since the leaves of the trees are, by default, mixing, for any fixed $\lambda \in (0, 1)$ and $h = 1$ it holds that

$$Q_{\{h, \lambda h\}} \leq \Pr[\deg(r(T)) \geq \Delta_+] = \sum_{t \geq \Delta_+} \xi_t \leq \exp [-2\mathcal{C}] \leq \exp [-(1 - \theta' \cdot \lambda) \mathcal{C}],$$

as $\lambda, \theta > 0$ while $\lambda \cdot \theta' < 1$. The lemma follows.

12 Proof of Proposition 1

Given some $\sigma_L \in [k]^L$, we let the variable $Y = Y(\sigma_L)$ be such that $Y = \mu_{r_T}^{\sigma_L}(c) - 1/k$. Let the colouring of the root $\tau_r = c$. By definition, we have that

$$\begin{aligned}\mathbb{E}_{\mu^{\tau_r}}[Y] &= \sum_{\sigma_L \in [k]^L} \mu_L^{\tau_r}(\sigma_L) Y(\sigma_L) \\ &= \sum_{\sigma_L \in [k]^L} \mu_L^{\tau_r}(\sigma_L) (\mu^{\sigma_L}(c) - 1/k) = \mu^{X(L)}(c) - 1/k.\end{aligned}$$

Also, we have that

$$\begin{aligned}\mathbb{E}_{\mu^{\tau_r}}[Y] &= \sum_{\sigma_L \in [k]^L} \frac{\mu_L^{\tau_r}(\sigma_L)}{\mu_L(\sigma_L)} (\mu^{\sigma_L}(c) - 1/k) \cdot \mu_L(\sigma_L) \\ &= \sum_{\sigma_L \in [k]^L} \frac{\mu_r^{\sigma_L}(c)}{\mu_r(c)} (\mu^{\sigma_L}(c) - 1/k) \cdot \mu_L(\sigma_L).\end{aligned}$$

That is, in order to compute the expectation above we calculate the Randon-Nikodym derivative. The derivation in the second line is just an application of Bayes' rule. Letting $\frac{\mu_r^{\sigma_L}(c)}{\mu_r(c)} = r(\sigma_L)$ and noting that $\mu_r(c) = 1/k$, it is elementary to verify that

$$k \cdot Y(\sigma_L) + 1 = r(\sigma_L).$$

Using the above equality we get that

$$\mathbb{E}_{\mu^{\tau_r}}[Y] = k \sum_{\sigma_L \in [k]^L} (\mu^{\sigma_L}(c) - 1/k)^2 \mu(\sigma_L) + \sum_{\sigma_L \in [k]^L} (\mu^{\sigma_L}(c) - 1/k) \mu(\sigma_L). \quad (32)$$

It is direct to show that $\sum_{\sigma_L \in [k]^L} (\mu^{\sigma_L}(c) - 1/k) \mu(\sigma_L) = 0$. Thus, we get that

$$\mathbb{E}_{\mu^{\tau_r}}[Y] = \mathbb{E}[Y^2] = \mu^{X(L)}(c) - 1/k. \quad (33)$$

where the second expectation is w.r.t. the unconditional Gibbs distribution. Observe that $\mathbb{E}_{\mu^{\tau_r}}[Y] \geq 0$.

Using the above equality and Cauchy-Schwarz inequality we get the following:

$$\begin{aligned}\sum_{\sigma(L) \in [k]^L} \mu_L(\sigma_L) \cdot \left| \mu_{r(T)}^{\sigma_L}(c) - 1/k \right| &\leq \sqrt{\sum_{\sigma(L) \in [k]^L} \mu_L(\sigma_L) \cdot \left| \mu_{r(T)}^{\sigma_L}(c) - 1/k \right|^2} \quad [\text{Cauchy-Schwarz}] \\ &\leq \sqrt{\frac{1}{k} \left| \mu_{r(T)}^{X_L}(c) - 1/k \right|}. \quad [\text{from (33)}] \quad (34)\end{aligned}$$

Observe that in (34) the quantity inside the absolute value is always non-negative (e.g. from 33). Also, it holds that

$$\left| \mu_{r(T)}^{X_L}(c) - 1/k \right| \leq \|\mu^{X_L}(\cdot) - \mu(\cdot)\|_{r_T} = \|\mu^{X_L}(\cdot) - \mu^{Z_L}(\cdot)\|_{r_T}. \quad (35)$$

where Z is a random k -colouring of T . The equality, above, holds since the distributions μ_{r_T} and $\mu_{r_T}^{Z_L}$ are identical. For every $q \in [k]$ let Z^q denote a random colouring of T conditional that $r(T)$ is coloured

q. By the definition of total variation distance we get the following:

$$\begin{aligned}
\|\mu^{X_L}(\cdot) - \mu^{Z_L}(\cdot)\|_{r_T} &= \frac{1}{2} \sum_{c' \in [k]} |\mu_{r_T}^{X_L}(c') - \mu_{r_T}^{Z_L}(c')| \leq \frac{1}{2} \sum_{c' \in [k]} \left| \mu_{r_T}^{X_L}(c') - \frac{1}{k} \sum_{q \in [k]} \mu_{r_T}^{Z_L^q}(c') \right| \\
&\leq \frac{1}{k} \sum_{q \in [k]} \frac{1}{2} \sum_{c' \in [k]} |\mu_{r_T}^{X_L}(c') - \mu_{r_T}^{Z_L^q}(c')| \\
&\leq \frac{1}{k} \sum_{q \in [k]} \left\| \mu^{X_L}(\cdot) - \mu^{Z_L^q}(\cdot) \right\|.
\end{aligned} \tag{36}$$

Since the r.h.s. of (36) is a convex combination, it follows that

$$\|\mu^{X_L}(\cdot) - \mu^{Z_L}(\cdot)\|_{r_T} \leq \max_{q \in [k]} \left\{ \left\| \mu^{X_L}(\cdot) - \mu^{Z_L^q}(\cdot) \right\| \right\}.$$

The proposition follows by combining the above inequality, (35) and (34).

13 Proof of Theorem 2 - Reconstruction

Consider the following.

Definition 12 (Freezable Root) Consider Δ_- and δ as in the statement of Theorem 2. For a tree T of height t , its root is freezable if the following holds: If $t = 1$, then $r(T)$ is of degree is at least Δ_- . If $t > 1$, $r(T)$ is freezable if and only if $\deg(r_T) \geq \Delta_-$ and there are at least $\Delta_- - (\Delta_-)^\delta$ many vertices v children of $r(T)$ such that \hat{T}_v has a freezable root.

Definition 13 (Freezing Boundary) Let T be a tree of height t , for some integer $t > 0$, and let $L = L_t(T)$. Let σ be a k -colourings of T , for some $k > 0$. Then the boundary condition σ_L freezes the colouring r_T if the following holds: There exists $c \in [k]$ such that $\mu_{r_T}^{\sigma_L}(c) = 1$.

That is, a freezing boundary condition forces a unique colouring assignment at the root T .

Let \mathcal{F}_h denote the set of trees of height h which have freezable root. Since the total variation distance is always non-negative, it holds that

$$\mathbb{E} \|\mu^i - \mu^j\|_{L_h} \geq \Pr \left[\mathcal{T}_\xi^h \in \mathcal{F}_h \right] \cdot \mathbb{E} \left[\|\mu^i - \mu^j\|_{L_h} \mid \mathcal{T}_\xi^h \in \mathcal{F}_h \right] \tag{37}$$

The proof is going to be done in two steps. We are going to show that taking $k = (1 - \alpha)\Delta_- / \ln \Delta_-$, both $\Pr \left[\mathcal{T}_\xi^h \in \mathcal{F}_h \right]$ and $\mathbb{E} \left[\|\mu^i - \mu^j\|_{L_h} \mid \mathcal{T}_\xi^h \in \mathcal{F}_h \right]$ are bounded away from zero, for any $h > 0$. In particular we have the following:

Lemma 5 Given ξ, δ, Δ_- as in Theorem 2 the following is true: It holds that $\Pr \left[\mathcal{T}_\xi^h \in \mathcal{F}_h \right] \geq 1 - g$, where g is from Definition 7.

Remark 3 Given ξ and Δ_- , we choose g to be the smallest number which satisfies (4). We should note that the quantity g does not depend on h , the height of the tree.

Proof of Lemma 5 We are going to use induction to show that $\Pr \left[\mathcal{T}_\xi^h \notin \mathcal{F}_h \right] < g$. For $h = 1$, we use Definition 12, i.e.

$$\Pr \left[\mathcal{T}_\xi^h \notin \mathcal{F}_h \right] = \Pr[\deg(r(\mathcal{T}_\xi^h)) < \Delta_-] = \sum_{i < \Delta_-} \xi_i \leq g,$$

where the last inequality follows from the definition of the quantity g , i.e. from Definition 7. Assume now that $\gamma = \Pr \left[\mathcal{T}_\xi^{h-1} \notin \mathcal{F}_{h-1} \right] \leq g$ is true for some $h > 1$. We are going to show that it is also true that $\Pr \left[\mathcal{T}_\xi^h \notin \mathcal{F}_h \right] \leq g$. Let the \mathcal{Y}_r denote the event that r_T has less than $(\Delta_-) - (\Delta_-)^\delta$ children which v such that \tilde{T}_v does not have a freezable root. It holds that

$$\begin{aligned}
\Pr \left[\mathcal{T}_\xi^h \notin \mathcal{F}_h \right] &\leq \Pr \left[\deg(r(\mathcal{T}_\xi^h)) < \Delta_- \right] + \Pr \left[\deg(r(\mathcal{T}_\xi^h)) \geq \Delta_- \right] \Pr[\mathcal{Y}_r | \deg(r(\mathcal{T}_\xi^h)) \geq \Delta_-] \\
&\leq \sum_{i < \Delta_-} \xi_i + \sum_{i \geq \Delta_-} \Pr[\mathcal{Y}_r, \deg(r(\mathcal{T}_\xi^h)) = i] \\
&\leq \sum_{i < \Delta_-} \xi_i + \sum_{i \geq \Delta_-} \xi_i \Pr \left[\mathcal{B}(i, 1 - \gamma) < (\Delta_-) - (\Delta_-)^\delta \right] \\
&\leq \sum_{i < \Delta_-} \xi_i + \sum_{i \geq \Delta_-} \xi_i \Pr \left[\mathcal{B}(i, 1 - g) < (\Delta_-) - (\Delta_-)^\delta \right] \leq g. \quad [\text{by Definition 7}]
\end{aligned}$$

The lemma follows. \square

Lemma 6 *Let α, δ, Δ_- be as in Theorem 2. For $k = (1 + \alpha)\Delta_- / \ln \Delta_-$ it holds that*

$$\mathbb{E} \left[\|\mu^i - \mu^j\|_{L_h} \mid \mathcal{T}_\xi^h \in \mathcal{F}_h \right] \geq \left(1 - \frac{2}{\log k} \right).$$

Proof: The lemma will follow by assuming any instance of the trees in \mathcal{F}_h , i.e. we consider a fixed tree $T \in \mathcal{F}_h$. We let \mathbf{F} denote the set of these vertices v children of $r(T)$ such that \tilde{T}_v has a freezable root. Since we have assumed that $T \in \mathcal{F}_h$ it holds that $|\mathbf{F}| \geq \Delta_- - (\Delta_-)^\delta$.

Take a random colouring of T . W.l.o.g. assume that the root is coloured with colour c . This means that each of the children of the root has a colour which is distributed uniformly at random in $[k] \setminus \{c\}$ and each of the colour assignments is independent of the other. So as the colour assignment of the root to be frozen, it suffices to have the following: For every colour $q \in [k] \setminus \{c\}$ there should be at least one child in \mathbf{F} which is assigned q and its colouring is frozen. Clearly, examining only the children of the $r(T)$ which are in \mathbf{F} will yield a lower bound for the probability that we have a frozen colouring at $r(T)$. Let P_h denote the probability that the root of T is frozen. For the Gibbs distribution of the tree T then it holds that

$$\|\mu^i - \mu^j\|_{L_h} \geq P_h.$$

Also, since the tree T is chosen arbitrarily from \mathcal{F}_h , we get that P_h is a lower bound for the expectation $\mathbb{E} \left[\|\mu^i - \mu^j\|_{L_h} \mid \mathcal{T}_\xi^h \in \mathcal{F}_h \right]$, too. The lemma follows by bounding appropriately P_h .

At this point, we can derive the bound by working, essentially, as in [27, 29, 30]. For the sake of completeness in what follows we present the steps for bounding P_h .

Letting w_q denote the number of occurrences of the colour q between the vertices in \mathbb{F} we have that

$$P_h = \mathbb{E} \left[\prod_{q \in [k] \setminus \{c\}} (1 - (1 - P_{h-1})^{w_q}) \right], \quad (38)$$

where the expectation is w.r.t. the random variables w_q . Clearly the variables w_q for different q follow the multinomial distribution. E.g. they should sum to $|\mathbf{F}|$. Clearly the random variables are correlated with each other.

Consider a set of $k - 1$ independent random variables \tilde{w}_q for every $q \in [k] \setminus \{c\}$. Each \tilde{w}_q follows a Poisson distribution with parameter $D = \frac{|\mathbf{F}|}{k-1} \left(1 - \frac{1}{\log k} \right)$. It is elementary to show that conditional that

$\sum_{q \in [k] \setminus \{c\}} \tilde{w}_q \leq |\mathbf{F}|$ there is a coupling of (w_1, \dots, w_{k-1}) and $(\tilde{w}_1, \dots, w_{k-1})$ such that for every q it holds that $w_q \geq \tilde{w}_q$, (e.g. see Lemma 4 in [30]). Then clearly we get that

$$\begin{aligned} P_h &\geq \mathbb{E} \left[\prod_{q \in [k] \setminus \{c\}} (1 - (1 - P_{h-1})^{\tilde{w}_q}) \right] - \Pr \left[\sum_{q \in [k] \setminus \{c\}} \tilde{w}_q > |\mathbf{F}| \right] \\ &\geq \prod_{q \in [k] \setminus \{c\}} \mathbb{E} [(1 - (1 - P_{h-1})^{\tilde{w}_q})] - \Pr \left[\sum_{q \in [k] \setminus \{c\}} \tilde{w}_q > |\mathbf{F}| \right] \\ &\geq [1 - \exp(P_{h-1}D)]^{k-1} - \Pr \left[\sum_{q \in [k] \setminus \{c\}} \tilde{w}_q > |\mathbf{F}| \right], \end{aligned}$$

in the second inequality we use the fact that \tilde{w}_q s are independent with each other. It holds that $\sum_{q \in [k] \setminus \{c\}} \tilde{w}_q$ is distributed as in $\text{Po}(|\mathbf{F}| (1 - 1/\log k))$. Thus, it holds that $s = \Pr \left[\sum_{q \in [k] \setminus \{c\}} \tilde{w}_q > |\mathbf{F}| \right] \leq 1/k^2$.

Let $f(x) = (1 - \exp(xD))^{k-1} - s$. Then it is direct to verify that $f(1 - \frac{1}{\log k}) > 1 - \frac{1}{\log k}$. Since $P_0 = 1$ and $f(x)$ is increasing function we get that $P_h > 1 - \frac{1}{\log k}$, for any $h \geq 0$. \square

14 Proof of Theorem 1

We will show the theorem by using Theorem 2.

Let ξ be a distribution on the non-negative integers such that it is well-concentrated. Also let d_ξ be the expected value of ξ . We assume that d_ξ is sufficiently large.

The theorem follows by showing that for any fixed $\alpha > 0$, for $k_1 = (1 + \alpha)d_\xi / \ln d_\xi$ and $k_2 = (1 - \alpha)d_\xi / \ln d_\xi$ the following is true: There exist appropriate numbers $\gamma_1 = \gamma_1(\alpha) > 0$ and $\gamma_2 = \gamma_2(\alpha) > 0$ such that $d_\xi \leq \Delta_+ \leq (1 + \gamma_1)d_\xi$ also $d_\xi \geq \Delta_- \geq (1 - \gamma_2)d_\xi$, where Δ_+ and Δ_- are chosen as specified by Theorem 2. Furthermore it holds that $k_1 \geq (1 + \alpha/2)\Delta_+ / \ln \Delta_+$ and $k_2 \leq (1 - \alpha/2)\Delta_- / \ln \Delta_-$.

Consider, first, the quantity Δ_+ . We choose γ_1 to be the largest number such that $(1 + \alpha)d_\xi / \ln d_\xi \geq (1 + \alpha/2)\rho / \ln \rho$, where $\rho = (1 + \gamma_1)d_\xi$. We choose γ_1 to be independent of d_ξ . This means that for a given α and γ_1 , the inequality $(1 + \alpha)d_\xi / \ln d_\xi \geq (1 + \alpha/2)\rho / \ln \rho$ holds for sufficiently large d_ξ .

It suffices to show that Δ_+ , chosen as specified in Theorem 2, is such that $d_\xi \leq \Delta_+ \leq (1 + \gamma_1)d_\xi$. Note that the parameter δ we use for Δ_+ is such that $\delta = \min\{\alpha/4, 1/10\}$.

Since ξ is well concentrated, for any $x \geq (1 + \gamma_1)d_\xi$ it holds that

$$\sum_{i \geq x} \xi_i \leq x^{-c}, \quad (39)$$

where $c > 0$ is sufficiently large number. Choosing $q = 2d_\xi^{-c}$ it is direct to verify that the condition (2) is trivially satisfied by choosing $\Delta_+ \leq (1 + \gamma_1)d_\xi$. This follows by using the inequality in (39), i.e. that ξ is well concentrated and the Chernoff bounds for $\Pr[\mathcal{B}(\Delta_+, q) \geq \Delta_+^\delta]$.

The leftmost conditions in (3) is also satisfied for $\Delta_+ \leq (1 + \gamma_1)d_\xi$ and sufficiently large $c > 0$. I.e. it holds that

$$\sum_{t > (1 + \gamma_1)d_\xi} t \cdot \xi_t \leq \sum_{t > (1 + \gamma_1)d_\xi} t \cdot t^{-c} \leq 2[(1 + \gamma_1)d_\xi]^{-(c-1)}.$$

The second condition in (3) is trivially satisfied, as we describe above.

Consider now the case of Δ_- . We work in a very similar way as for the case of Δ_+ . We choose γ_2 to be the largest number such that $(1 - \alpha)d_\xi / \ln d_\xi \leq (1 - \alpha/2)\rho / \ln \rho$, where $\rho = (1 - \gamma_2)d_\xi$. We choose γ_2 to be independent of d_ξ , in the same manner as we chose γ_1 , for Δ_+ .

It suffices to show that Δ_- , chosen as specified in Theorem 2, is such that $d_\xi \geq \Delta_- \leq (1 - \gamma_2)d_\xi$. Note that the parameter δ we use for Δ_- is such that $\delta = \min\{\alpha/4, 1/10\}$.

Our assumption that ξ is well concentrated, implies that

$$\sum_{i \leq (1-\gamma_2)d_\xi} \xi_i \leq d_\xi^{-c}. \quad (40)$$

Setting $d_\xi \geq \Delta_- \geq (1 - \gamma_2)d_\xi$ and $g = 2d_\xi^{-c}$, where c is the same as above, it suffices to show that the constraint (4), in Definition 7, is satisfied. In particular, in the light of (39), it suffices to show that for our choice of g and Δ_- , the rightmost sum in (4) is sufficiently small.

It holds that $g \cdot \Delta_- < d_\xi^{-c/2} \ll (\Delta_-)^{-1+\delta}$. This implies that for any $i \geq \Delta_-$ we have that

$$\Pr \left[\mathcal{B}(i, 1 - g) < (\Delta_-) - (\Delta_-)^\delta \right] < \Pr \left[\mathcal{B}(\Delta_-, 1 - g) < (\Delta_-) - (\Delta_-)^\delta \right],$$

as $\Delta_- - \Delta_-^\delta < i \cdot g$ for all $i \geq \Delta_-$. Thus, it holds that

$$\begin{aligned} \sum_{i \geq \Delta_-} \xi_i \Pr \left[\mathcal{B}(i, 1 - g) < (\Delta_-) - (\Delta_-)^\delta \right] &\leq \Pr \left[\mathcal{B}(\Delta_-, 1 - g) < (\Delta_-) - (\Delta_-)^\delta \right] \sum_{i \geq \Delta_-} \xi_i \\ &\leq \Pr \left[\mathcal{B}(\Delta_-, 1 - g) < (\Delta_-) - (\Delta_-)^\delta \right] \\ &= \Pr \left[\mathcal{B}(\Delta_-, g) > (\Delta_-)^\delta \right] \leq \exp \left(-\Delta_-^\delta \right). \end{aligned}$$

The inequality in the second line follows from the fact that $\sum_{i \geq \Delta_-} \xi_i \leq 1$. The last inequality follows from a direct application of Chernoff bounds, i.e. Corollary 2.4 in [19]. Using the above bounds, it is trivial to show for our choice of g and Δ_- (4) is true.

The theorem follows.

Acknowledgement. The author of this work would like to thank Guilhem Semerjian for our communication and the discussion on the problem. Also, the author would like to thank Amin Coja-Oghlan for the discussions, his comments and the suggestions for improving the content of this work.

References

- [1] D. Achlioptas, A. Coja-Oghlan. *Algorithmic Barriers from Phase Transitions*. In Proc. of 49th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2008.
- [2] N. Bhatnagar, J. Vera, E. Vigoda and D. Weitz. *Reconstruction for coloring on Trees*. SIAM J. Discrete Math. **25** (2), (2011), pp 809-826.
- [3] N. Bhatnagar, A. Sly, and P. Tetali. *Reconstruction Threshold for the Hardcore Model*. In Proc. of the 14th International Conference on Randomization and Computation (RANDOM), 6302 LNCS, pp 434-447, 2010.
- [4] C. Borgs, J. Chayes, E. Mossel, and S. Roch. *The kesten-stigum reconstruction bound is tight for roughly symmetric binary channels*. In Proc. of the 47th Annual IEEE Symposium on Foundations of Computer Science (FOCS06), pp 518530, 2006.
- [5] J. T. Chayes, L. Chayes, J. Sethna and D. J. Thoules. *A mean field spin glass with short-range interactions*. Comm. Math. Phys., **106**(1), pp41-89, 1986.

- [6] Amin Coja-Oghlan. *A better algorithm for random k -SAT*. SIAM Journal on Computing 39:2823-2864.
- [7] A. Coja-Oghlan C. Efthymiou. *On independent sets in random graphs*. To appear in the journal Random Structures and Algorithms. Also, in proc. of 22nd Symposium on Discrete Algorithms (SODA'11), pp 136-144, 2011.
- [8] A. Coja-Oghlan, C. Efthymiou, N. Jaafari. *Local convergence of random graph colorings*. Available in <http://arxiv.org/abs/1501.06301>
- [9] C. Daskalakis, E. Mossel and S. Roch. *Optimal Phylogenetic Reconstruction*. In proc. of the 38th Annual ACM Symposium on Theory of Computing (STOC), 159-168, 2006.
- [10] M. Dyer, A. Flaxman, A. M. Frieze and E. Vigoda. *Random colouring sparse random graphs with fewer colours than the maximum degree*. Random Struct. and Algorithms **29**, pp. 450-465, 2006.
- [11] C. Efthymiou, *MCMC sampling colourings and independent sets of $G(n, d/n)$ near uniqueness threshold*. In proc of SODA'14, pp 305-316, 2014.
- [12] C. Efthymiou *A simple algorithm for sampling colouring of $G(n, d/n)$ using $(2 + \epsilon)d$ colours*. In proc of SODA'12, pp 272-280, 2012.
- [13] C. Efthymiou *Switching Colouring of $G(n, d/n)$ for Sampling up to Gibbs Uniqueness Threshold*. To appear in ESA 2014.
- [14] W. Evan, C. Kenyon, Y. Peres and L. J. Schulman. *Broadcasting on trees, the Ising model*. The Annals of Applied Probability, 2000, Vol. 10 No. 2, , 410-433
- [15] A. Gerschenfeld and A. Montanari. *Reconstruction for models on random graphs*. In Proc. of 48th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2007.
- [16] H. O. Georgii. *Gibbs Measures and Phase Transitions*, de Gruyter Stud. Math. 9, de Gruyter, Berlin, 1988
- [17] G. R. Grimmett and C. J. H. McDiarmid. *On colouring random graphs*. Math. Proc. of the Camb. Phil. Soc. **77** (02), pp 313-32, 1975.
- [18] F. Krzakala, A. Montanari, F. Ricci-Tersenghi, G. Semerjian, L. Zdeborova. *Gibbs states and the set of solutions of random constraint satisfaction problems*. In Proc. National Academy of Sciences 104 (2007) pp 10318-10323.
- [19] S. Janson, T. Luczak and A. Ruciński. *Random Graphs*. John Wiley & Sons, Inc. 2000
- [20] F. Martinelli, A. Sinclair and D. Weitz. *Glauber dynamics on trees: Boundary conditions and mixing time*. Comm. Math. Phys., **250** (2004), pp 301-334.
- [21] F. Martinelli, A. Sinclair and D. Weitz. *Fast mixing for independent sets, colorings and other models on trees*. Random Structures Algorithms, **31**, (2007), pp 134-172.
- [22] M. Mézard, G. Parisi and R. Zecchina. *Analytic and Algorithmic Solution of Random Satisfiability Problems*. Science **297** no. 5582 pp. 812-815, 2002.
- [23] M. Molloy. *The freezing threshold for k -colourings of a random graph*. In proc. of the 44th ACM Symposium on Theory of Computing (STOC'12), pp 921-930, 2012.

- [24] A. Montanari, R. Restrepo and P. Tetali. *Reconstruction and Clustering in Random Constraint Satisfaction Problems*, SIAM Journal on Discrete Mathematics, 2011
- [25] E. Mossel. *Reconstruction on Trees: Beating the second eigenvalue*. The Annals of Applied Probability 2001, Vol. 11, No. 1 285-300.
- [26] E. Mossel. *Phase transitions in phylogeny*. Trans. Amer. Math. Soc., 356(6):2379-2404, 2004
- [27] E. Mossel and Y. Peres. *Information flow on trees*. The Annals of Applied Probability 2003, Vol. 13, No. 3, 817-844.
- [28] R. Motwani & P. Raghavan. *Randomized Algorithms*. Cambridge University Press, 1995.
- [29] G. Semerjian. *On the freezing of variables in random constraint satisfaction problems*. J. Stat. Phys, 2008, Vol. 130, No 2, 251-293.
- [30] A. Sly. *Reconstruction of Random Colourings*, Commun. Math. Phys., **188** (2009), pp 943-961.
- [31] Y. Yin, C. Zhang, *Sampling colorings almost uniformly in sparse random graphs*. In <http://arxiv.org/abs/1503.03351>